PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07H 21/00, C12Q 1/68 — A1

(11) International Publication Number: WO 95/01364

(43) International Publication Date: 12 January 1995 (12.01.95)

(21) International Application Number: PCT/US94/07234

(22) International Filing Date: 24 June 1994 (24.06.94)

(30) Priority Data:
083,088 25 June 1993 (25.06.93) US

(71) Applicant: YALE UNIVERSITY [US/US]; 451 College Street, New Haven, CT 06511 (US).

(72) Inventors: GLAZER, Peter, M.; 49 Thirsty Hill Road, Guilford, CT 06437 (US). HAVRE, Pamela, A.; 433 Whitney Avenue, Apartment 4B, New Haven, CT 06511 (US).

(74) Agents: PABST, Patrea, L. et al.; Kilpatrick & Cody, 1100 Peachtree Street, Suite 2800, Atlanta, GA 30309-4530 (US).

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: CHEMICALLY MODIFIED OLIGONUCLEOTIDE FOR SITE-DIRECTED MUTAGENESIS

(57) Abstract

A mutagenic, triplex-forming oligonucleotide and methods for use thereof wherein the oligonucleotide is chemically modified to incorporate a mutagen and forms a triple-stranded nucleic acid molecule with a specific DNA segment of a target DNA molecule. Upon formation of the triplex, the mutagen is brought into proximity with the target molecule and causes a mutation at a specific site therein. The mutation activates, inactivates or alters the activity and function of the target molecule.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgystan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

CHEMICALLY MODIFIED OLIGONUCLEOTIDE FOR SITE-DIRECTED MUTAGENESIS

Background of the Invention

This relates to the fields of genetics, and more particularly relates to site-directed mutagenesis of a gene of interest.

Gene Therapy

Gene therapy is the introduction into a cell of an entire replacement copy of a defective gene to treat human, animal and plant genetic disorders. The introduced gene, via genetic recombination, replaces the endogenous gene. This approach requires complex delivery systems to introduce the replacement gene into the cell, such as genetically engineered viruses, or viral vectors.

Gene therapy is being used on an experimental basis to treat well known genetic disorders of humans such as retinoblastoma, cystic fibrosis, and sickle cell anemia. However, *in vivo* efficiency is low due to the limited number of recombination events actually resulting in replacement of the defective gene.

Triple-stranded DNA

Since the initial observation of triple-stranded DNA many years ago by Felsenfeld et al., *J. Am. Chem. Soc.* 79:2023 (1957), oligonucleotide-directed triple helix formation has emerged as a valuable tool in molecular biology. Current knowledge suggests that oligonucleotides can bind as third strands of DNA in a sequence specific manner in the major groove in polypurine/polypyrimidine stretches in duplex DNA. In one motif, a polypyrimidine oligonucleotide binds in a direction parallel to the purine strand in the duplex, as described by Moser and Dervan, *Science* 238:645 (1987), Praseuth

et al., *Proc. Natl. Acad. Sci. USA* 85:1349 (1988), and Mergny et al., *Biochemistry* 30:9791 (1991). In the alternate purine motif, a polypurine strand binds anti-parallel to the purine strand, as described by Beal and Dervan, *Science* 251:1360 (1991). The specificity of triplex formation arises from base triplets (AAT and GGC in the purine motif) formed by hydrogen bonding; mismatches destabilize the triple helix, as described by Mergny et al., *Biochemistry* 30:9791 (1991) and Beal and Dervan, *Nuc. Acids Res.* 11:2773 (1992).

Triplex forming oligonucleotides have been found useful for several molecular biology techniques. For example, triplex forming oligonucleotides designed to bind to sites in gene promoters have been used to block DNA binding proteins and to block transcription both *in vitro* and *in vivo*. (Maher et al., *Science* 245:725 (1989), Orson et al., *Nucleic Acids Res.* 19:3435 (1991), Postal et al., *Proc. Natl. Acad. Sci. USA* 88:8227 (1991), Cooney et al., *Science* 241:456 (1988), Young et al., *Proc. Natl. Acad. Sci. USA* 88:10023 (1991), Maher et al., *Biochemistry* 31:70 (1992), Duval-Valentin et al., Proc. Natl. Acad. Sci. USA 89:504 (1992), Blume et al., *Nucleic Acids Res.* 20:1777 (1992), Durland et al., *Biochemistry* 30:9246 (1991), Grigoriev et al., *J. of Biological Chem.* 267:3389 (1992), and Takasugi et al., *Proc. Natl. Acad. Sci. USA* 88:5602 (1991)). Site specific cleavage of DNA has been achieved by using triplex forming oligonucleotides linked to reactive moieties such as EDTA-Fe(II) or by using triplex forming oligonucleotides in conjunction with DNA modifying enzymes (Perrouault et al., *Nature* 344:358 (1990), Francois et al., *Proc. Natl. Acad. Sci. USA* 86:9702 (1989), Lin et al., *Biochemistry*

28:1054 (1989), Pei et al., *Proc. Natl. Acad. Sci. USA* 87:9858 (1990), Strobel et al., *Sci nce* 254:1639 (1991), and Posvic and Dervan, *J. Am. Chem Soc.* 112:9428 (1992)). Sequence specific DNA purification using triplex affinity capture has also been demonstrated. (Ito et al., *Proc. Natl. Acad. Sci. USA* 89:495 (1992)). Triplex forming oligonucleotides linked to intercalating agents such as acridine, or to cross-linking agents, such as p-azidophenacyl and psoralen, have been utilized, but only to enhance the stability of triplex binding. (Praseuth et al., *Proc. Natl. Acad. Sci. USA* 85:1349 (1988), Grigoriev et al., *J. of Biological Chem.* 267:3389 (1992), Takasugi et al., *Proc. Natl. Acad. Sci. USA* 88:5602 (1991).

A method for site-directed mutagenesis of a target DNA molecule would be a useful in achieving successful gene or anti-viral therapy. Such a method would also be a useful research tool for genetic engineering or for studying genetic mechanisms such as DNA repair.

Therefore, it is an object of the present invention to provide a method for *in vivo* and *in vitro* site-directed mutagenesis of a target DNA molecule.

It is a further object of the present invention to provide a method for mutagenesis of a target DNA molecule that is highly specific and efficient.

It is a further object of the present invention to provide a method for treating genetic disorders by gene therapy without the need for a viral vector.

It is a further object of the present invention to provide a method for treating cancer.

It is a further object of the present invention to provide a mutagenic oligonucleotide for use in therapy and research.

Summary of th Invention

A mutagenic, triplex-forming oligonucleotide and methods for use thereof are described herein. An oligonucleotide capable of forming a triple strand with a specific DNA segment of a target gene DNA is chemically modified to incorporate a mutagen. The modified oligonucleotide hybridizes to a chosen site in the target gene, forming a triplex region, thereby bringing the attached mutagen into proximity with the target gene and causing a mutation at a specific site in the gene. The mutation activates, inactivates, or alters the activity and function of the target gene.

If the target gene contains a mutation that is the cause of a genetic disorder, then the mutagenic oligonucleotide is useful for mutagenic repair that may restore the DNA sequence of the target gene to normal. If the target gene is a viral gene needed for viral survival or reproduction or an oncogene causing unregulated proliferation, such as in a cancer cell, then the mutagenic oligonucleotide is useful for causing a mutation that inactivates the gene to incapacitate or prevent reproduction of the virus or to terminate or reduce the uncontrolled proliferation of the cancer cell. The mutagenic oligonucleotide is also a useful anti-cancer agent for activating a repressor gene that has lost its ability to repress proliferation.

The mutagenic triplex-forming oligonucleotide is also particularly useful as a molecular biology research tool to cause site-directed or targeted mutagenesis. Site-directed mutagenesis is useful for targeting a normal gene and for the study of mechanisms such as DNA repair. Targeted mutagenesis of a specific gene in an animal oocyte, such as a mouse oocyte, provides a useful and

powerful tool for genetic engineering for research and therapy and for generation of new strains of "transmutated" animals and plants for research and agriculture.

Brief Description of the Drawings

Figure 1 is a schematic representation showing a psoralen-linked triplex-forming oligonucleotide for targeted mutagenesis of the lambda genome. A map of the lambda supF genome is shown, including the target gene for site-directed mutagenesis, the supF suppressor tRNA gene. Above the partial sequence of the supF gene (positions 149 to 183), the site of triplex formation at positions 167-176 is indicated by the placement of the triplex-forming oligonucleotide, pso-AG10 (4'hydroxymethyl-4,5',8-trimethylpsoralen-$^{5'}$AGGAAGGGGG$^{3'}$). The arrow indicates that the psoralen moiety is targeted to the A:T base pair at position 167. In addition to the supF gene, the lambda vector carries the cI lambda repressor gene which is used to assess non-targeted mutagenesis.

Figures 2a and 2b show a sequence analysis of targeted mutagenesis in the supF gene by the psoralen-linked triplex-forming oligonucleotide (pso-AG10). In Figure 2a, mutations produced by pso-AG10 and UVA are indicated above each base pair, with the listed base representing the change from the sequence in the top strand. The + signs below the sequence are sites at which mutations are known to produce a detectable phenotype change, demonstrating that the use of supF in this assay does not bias detection at any particular site. The asterisk indicates the targeted base pair at position 167. DNA sequence data was obtained by automated methods after polymerase chain reaction

amplification of the supF genes from lambda phage plaques in accordance with the method of Connell et al., *Biotechniques* 5:342 (1987). Figure 2b is a compilation of mutations induced in supF by 8-methoxypsoralen and UVA in mouse L-cells using the lambda supF vector of Figure 1 or generated using a plasmid shuttle vector in monkey Vero cells in accordance with the method of Bredberg and Nachmansson, *Carcinogenesis* 8:1923 (1987), to compare the mutations produced in supF by free psoralen with those produced by the triplex forming oligonucleotide AG10.

Figure 3 is a schematic representation of the strategy for targeted mutagenesis of SV40 DNA. The 10 base triplex-forming oligonucleotide, psoralen-AG10 (4'hydroxymethyl-4,5',8-trimethyl-5'AGGAAGGGGG3'), is shown directly above its targeted sequence in the supF gene (base pairs 167-176), contained within the SV40 vector, pSP189. Psoralen-AG10 is incubated with the SV40 vector DNA to allow site-specific triplex formation. Photoactivation of the psoralen by irradiation with long wave ultraviolet light (320-400 nm) is designed to generate an adduct at the targeted base pair (167), as indicated by the arrow. The oligo-plasmid complex is then transfected into monkey COS-7 cells and allowed to replicate for 48 hours. Following purification of the vector DNA by the Hirt lysate procedure (Hirt et al., *J. Mol. Biol.* 26:365-369 (1967)), the DNA is used to transform E. coli SY204 *lacZ*125 (Am). Transformants are selected on ampicillin plates containing X-gal and isopropylthio-ß-D-galactoside (IPTG) for detection and isolation of mutants (white colonies) in which the supF gene has been inactivated by mutation.

Figure 4 is a schematic representation showing the basis of a restriction enzyme protection assay,

using *Hinf* I digestion, to detect site-specific triplex formation within the supF gene. The formation of triplex DNA by psoralen-AG10 at its targeted site (bp 167-176 of the supF gene) overlaps with the *Hinf* I restriction site at bp 164-168 (indicated in the diagram by the box around the appropriate base pairs). Digestion of the unprotected 250 bp supF PCR fragment with *Hinf* I is expected to yield three fragments of sizes 150, 65, and 35 bp. In contrast, with the *Hinf* I site at bp 164-168 blocked by triplex formation at bp 167-176, fragments of sizes 150 and 100 bp are predicted.

Figures 5A and 5B are a sequence analysis of targeted mutagenesis in the supF gene within the pSP189 SV40 vector by the psoralen-linked triplex-forming oligonucleotide, psoralen-AG10. In Figure 5A, point mutations produced by psoralen-AG10 and UVA are indicated above each base pair, with the listed base representing the change from the sequence in the top strand. Deletion mutations are presented below the supF sequence, indicated by dashed lines. For the one deletion that was accompanied by an apparent base change, the indicated base represents a mutation from the sequence of the top strand. The + signs below the sequence are sites at which mutations are known to produce a detectable phenotype change demonstrating that the use of supF in this assay does not bias detection at any particular site. The asterisk indicates the targeted base pair at position 167. Figure 5B is a compilation of mutations induced in supF by 8-methoxypsoralen and UVA in mouse L-cells using a lambda phage shuttle vector or generated in monkey Vero cells using an SV40 shuttle vector (pZ189) almost identical to the one used in this study to show for comparison the mutations that can be produced in supF by free psoralen.

Figure 6 is a schematic representation of the strategy for targeted mutagenesis of SV40 DNA in monkey COS c 11s. Psoralen was incorporated by synthesis into a triplex-forming oligonucleotide as psoralen phosphoramidite. The SV40 shuttle vector, pSupFG1 was a derivative of pSP189 and carried triplex-binding sites, which were engineered in to the *supF* gene. The plasmid is transfected into monkey COS-7 cells. Subsequently, the oligonucleotide is added to the cells, which are allowed to replicate for 48 hours. Photoactivation of the psoralen by irradiation with long wave ultraviolet light (320-400 nm) is designed to generate mutations. Following purification of the vector DNA by the Hirt lysate procedure (Hirt et al., *J. Mol. Biol.* 26:365-369 (1967)), the DNA is used to transform E. coli SY204 *lacZ*125 (Am). Transformants are selected on ampicillin plates containing X-gal and IPTG for detection and isolation of mutants (white colonies) in which the supF gene has been inactivated by mutation.

Figure 7 is the nucleotide sequence and two-dimensional structure of a modified *supF* gene, designated *supFG1a*. An A:T to C:G transversion was incorporated into the sequence, along with a compensatory T:A to G:C change at bp 101 to maintain base pairing in the amino acid acceptor stem of the mature tRNA thereby eliminating an interruption at bp 167 in the polypurine/polypyrimidine run. In addition, a 13 bp polypurine/polypyrimidine sequence was inserted between bp 183 and 184 to extend the length of the polypurine/polypyrimidine run in the gene to 30 bp. This construct contains a 30 bp polypurine site with two interruptions.

Figure 8 is the nucleotide sequence and two-dimensional structure of a modified *supF* gene,

designated *supFG2*. This sequence contains a 43 bp polypurine site and two interruptions.

Figure 9 is a graph of percent maximum binding versus oligonucleotide concentration (M) of the mutagenic oligonucleotides pso-AG10 (open squares) and pso-AGT30 (diamonds).

Figure 10 is a sequence analysis of sequences of mutations targeted to the supF gene within an SV40 vector (pSupFG1a) by treatment of COS cells with pso-AGT30 and UVA.

Detailed Description of the Invention

A mutagenic triplex-forming oligonucleotide and methods of use in gene therapy, anti-viral therapeutics, scientific research, and genetic engineering of cells, animals and plants are provided. The mutagenic oligonucleotide binds with specificity to a chosen site in a target DNA molecule, forming a triplex region, thereby bringing the attached mutagen into proximity with the target site and causing a mutation therein. Preferably, the mutation activates, inactivates or alters the activity and function of a gene containing the target site.

Oligonucleotide

The oligonucleotide is a synthetic or isolated oligonucleotide capable of binding or hybridizing with specificity to a predetermined region of a double-stranded DNA molecule to form a triple-stranded structure. Preferably, the predetermined region of the double-stranded molecule contains or is adjacent to the defective or essential portion of a target gene, such as the site of a mutation causing a genetic defect, a site causing oncogene activation, or a site causing the inhibition or

inactivation of an oncogene suppressor. Most preferably, the gene is a human gene.

Preferably, the oligonucleotide is a single-stranded DNA molecule between 7 and 40 nucleotides in length, most preferably 10 to 20 nucleotides in length for *in vitro* mutagenesis and 20 to 30 nucleotides in length for *in vivo* mutagenesis. The base composition is preferably homopurine or homopyrimidine. Alternatively, the base composition is polypurine or polypyrimidine. However, other compositions are also useful. The preferred conditions under which a triple-stranded structure will form and the desired nucleotide composition of the third strand are well known to those skilled in the art. (See for example, Moser and Dervan, *Science* 238:645 (1987); Praseuth et al., *Proc. Natl. Acad. Sci. USA* 85:1349 (1988); Mergny et al., *Biochemistry* 30:9791 (1991); Beal and Dervan, *Science* 251:1360 (1991); Mergny et al., *Biochemistry* 30:9791 (1991) and Beal and Dervan, *Nuc. Acids Res.* 11:2773 (1992), which are incorporated by reference herein.)

Preferably, the mutagenic oligonucleotide hybridizes to the target nucleic acid molecule under conditions of high stringency and specificity. Most preferably, the oligonucleotide binds in a sequence-specific manner in the major groove of duplex DNA. Reaction conditions for *in vitro* triple helix formation of an oligonucleotide probe or primer to a nucleic acid sequence vary from oligonucleotide to oligonucleotide, depending on factors such as oligonucleotide length, the number of G:C and A:T base pairs, and the composition of the buffer utilized in the hybridization reaction. A mutagenic oligonucleotide substantially complementary, based on the third strand binding code, to the target

region of the double-stranded nucleic acid molecul is preferred.

A useful measure of triple helix formation is the equilibrium dissociation constant, K_d, of the triplex, which can be estimated as the concentration of mutagenic oligonucleotide at which triplex formation is half-maximal. Preferably, the oligonucleotide has a binding affinity for the target sequence in the range of physiologic interactions. The preferred mutagenic oligonucleotide has a K_d less than or equal to approximately 8×10^{-7} M. Most preferably, the K_d is less than or equal to 8×10^{-9} M in order to achieve significant intracellular interactions.

Mutagen

The oligonucleotide is chemically modified to include a mutagen at either the 5′ end, 3′ end, or internal portion so that the mutagen is proximal to the site in the gene requiring modification. Preferably, the mutagen is incorporated into the oligonucleotide during nucleotide synthesis. For example, commercially available compounds such as psoralen C2 phosphoramidite (Glen Research, Sterling, VA) are inserted into a specific location within an oligonucleotide sequence in accordance with the methods of Takasugi et al., *Proc. Natl. Acad. Sci. U.S.A.* 88:5602-5606 (1991), Gia et al., *Biochemistry* 31:11818-11822 (1992), Giovannangeli et al., *Nucleic Acids Res.* 20:4275-4281 (1992) and Giovannangeli et al., *Proc. Natl. Acad. Sci. U.S.A.* 89:8631-8635 (1992), all of which are incorporated by reference herein.

The mutagen may also be attached to the oligonucleotide by a covalent bond. For example, the mutagen is attached to the oligonucleotide by a linker, such as sulfo-*m*-maleimidobenzoly-*N*-hydroxysuccinimide ester (sulfo-MBS, Pierce

Chemical Co., Rockford, IL) in accordance with the methods of Liu et al., *Biochem.* 18:690-697 (1979) and Kitagawa and Ailawa, *J. Biochem.* 79:233-236 (1976), both of which are incorporated by reference herein. Alternatively, the mutagen is attached to the oligonucleotide by photoactivation, which causes the mutagen, such as psoralen, to bind to the oligonucleotide.

The mutagen can be any chemical capable of causing a mutation at the desired site of the double-stranded DNA molecule. Preferably the mutation restores the normal, functional sequence of the gene, inactivates an oncogene or activates an oncogene suppressor, or alters the function or inactivates a viral gene.

The chemical mutagen can either cause the mutation spontaneously or subsequent to activation of the mutagen, such as, for example, by exposure to light.

Preferred mutagens include psoralen, which requires activation by UVA irradiation, acridine orange, which can be activated by UVA irradiation and can be effective in the absence of light, and alkylating agents, cis-platinum analogs, hematoporphyrins and hematoporphyrin derivatives, mitomycin C, radionuclides such as ^{125}I, ^{35}S and ^{32}P, and molecules that interact with radiation to become mutagenic, such as boron that interacts with neutron capture and iodine that interacts with auger electrons. In particular, acridine orange can be used to cause a frame shift mutation, useful for gene inactivation.

If necessary for activation of the mutagen, light can be delivered to cells on the surface of the body, such as skin cells, by exposure of the area requiring treatment to a conventional light source. Light can be delivered to cells within the

body by fiber optics or laser by methods known to those skilled in the art. Targeted fluorogens that provide sufficient light to activate the light-activated mutagens can also provide a useful light source.

Method of Administration

Preferably, the mutagenic oligonucleotides are dissolved in a physiologically-acceptable carrier, such as an aqueous solution or are incorporated within liposomes, and the carrier or liposomes are injected into the organism undergoing genetic manipulation, such as an animal requiring gene therapy or anti-viral therapeutics. The preferred route of injection in mammals is intravenous. It will be understood by those skilled in the art that oligonucleotides are taken up by cells and tissues in animals such as mice without special delivery methods, vehicles or solutions.

For *in vitro* research studies, a solution containing the mutagenic oligonucleotides is added directly to a solution containing the DNA molecules of interest in accordance with methods well known to those skilled in the art and described in more detail in the examples below. *In vivo* research studies are conducted by transfecting cells with plasmid DNA and incubating the mutagenic oligonucleotides in a solution such as growth medium with the transfected cells for a sufficient amount of time for entry of the oligomers into the cells and for triplex formation. The transfected cells may be in suspension or a monolayer attached to a solid phase, or may be cells within a tissue wherein the oligonucleotide is in the extracellular fluid. The cells are then irradiated to activate the psoralen to form photoadducts and consequently mutations at the targeted site.

Methods of Use

If the target gene contains a mutation that is the cause of a genetic disorder, then the mutagenic oligonucleotide is useful for mutagenic repair that may restore the DNA sequence of the target gene to normal. For example, the mutagenic oligonucleotide may be useful for mutagenic repair of a defective gene such as the human ß-hemoglobin gene in sickle cell anemia, thalassemia and other hemoglobinopathies. If the target gene is an oncogene causing unregulated proliferation, such as in a cancer cell, then the mutagenic oligonucleotide is useful for causing a mutation that inactivates the gene and terminates or reduces the uncontrolled proliferation of the cell. The mutagenic oligonucleotide is also a useful anti-cancer agent for activating a repressor gene that has lost its ability to repress proliferation. Furthermore, the mutagenic oligonucleotide is useful as an antiviral agent when the oligonucleotide is specific for a portion of a viral genome necessary for proper proliferation or function of the virus.

The mutagenic triplex-forming oligonucleotide can also be used as a molecular biology research tool to cause site-directed mutagenesis in any gene for the study of mechanisms such as, for example, DNA repair. The oligonucleotide may also be used to study DNA repair by delivering an adduct to the DNA and studying how the adduct is processed into a mutation under various experimental conditions.

The mutagenic triplex-forming oligonucleotides will be further understood in view of the following non-limiting examples.

Example 1: Site-specific, Targeted Mutagenesis of the *supF* gene of the Lambda Phage Genome

A triplex-forming oligonucleotide linked to psoralen at its 5′ end was used to achieve site-specific, targeted mutagenesis in a specific gene in an intact, double-stranded lambda phage genome. Psoralen-linked oligonucleotides were obtained from either Oligos Etc. (Wilsonville, OR) or M. Talmor (Yale University, New Haven, CT) with materials from Glen Research (Sterling, VA). The psoralen was incorporated in the oligonucleotide synthesis as a psoralen phosphoramidite in accordance with the instructions provided by supplier.

As shown schematically in Figure 3, site-specific triplex formation was designed to deliver the psoralen to the targeted site in the lambda DNA, and UVA irradiation was used to activate the psoralen to form adducts and thereby induce mutations at that site. Sequence analysis of mutations in the target gene showed that almost all were in the targeted region, and 56% were found to be the same T:A to A:T transversion at the targeted base pair. The ratio of targeted to non-targeted mutagenesis was estimated by simultaneous analysis of mutagenesis in a non-targeted gene within the lambda genome, along with analysis of mutagenesis induced by a non-triplex forming (but psoralen linked) oligonucleotide. It was found that targeted mutations were produced at a frequency at least 500-fold greater than that of non-targeted mutations.

The target gene chosen was supF, an E. coli amber suppressor tyrosine tRNA gene, contained within the genome of a lambda phage vector, lambda supF as shown in Fig. 1. A 10 base homopurine oligonucleotide AG10 (5′ AGGAAGGGGG 3′) capable of forming a triple strand at positions 167-176 in the supF gene was identified. The ability of AG10 to

bind to the supF gene was demonstrated using ^{32}P-labeled AG10 in an in vitro binding reaction with a 250 bp fragment containing the entire supF gene.

To demonstrate targeted, site-specific triplex formation as a prelude to mutagenesis studies, binding assays were carried out for 2 hours at 37 degrees in 10% sucrose, 20 mM $MgCl_2$, 10 mM Tris (pH 8.0), and 1 mM spermidine in a 10 μl volume. The 250 bp supF target was generated from lambda supF using the polymerase chain reaction. Each oligo (200 ng) was labelled with 50 μCi of gamma-^{32}P-ATP (Amersham, Arlington Heights, IL) and separated from unreacted gamma^{32}P-ATP by passage through a G-25 spin column (Boehringer Mannheim, Indianapolis, IN). The concentration of oligomer in the reaction mixture was 6 x 10^{-8} M and the oligomer:supF ratio was approximately 1:1 on a molar basis. When present, competitors were used at 200-fold molar excess.

Following the 2 hour binding step, reaction mixtures were run on a 4% acrylamide gel in 90 mM Tris base, 90 mM boric acid, 20 mM $MgCl_2$ with a 20% acrylamide plug. A 100 bp ladder (BRL, Bethesda, MD) was end-labelled as described for oligomers and run on gels as a size reference. Following a 4 hour run at constant voltage (150 V), the gel was visualized by autoradiography for 1 hour using Kodak X-AR film. The electrophoretic gel showed binding of the triplex forming oligonucleotide "AG10" to the supF gene target. To assay for triplex formation, ^{32}P-labelled oligonucleotides, either AG10 ($^{5'}$AGGAAGGGGG$^{3'}$) or the reverse sequence oligomer (GA10), were incubated with a 240 bp double-stranded fragment containing the entire supF gene. The products of the binding reactions were visualized by polyacrylamide gel electrophoresis and autoradiography.

In the electrophoretic gel, binding of labelled AG10 to the added supF DNA (lane 2) was demonstrated by the new band migrating at the position appropriate to the 250 bp supF fragment. When no supF target DNA was present, there was no band observed at this position (lane 1). Excess unlabelled AG10 competed with the ^{32}P-labelled AG10 (lane 3), whereas an excess of the reverse sequence oligomer (GA10, 5′ GGGGGAAGGA 3′) did not compete with AG10 (lane 4). In lanes 5-8, no binding of the ^{32}P-labelled GA10 to supF was detected: (lane 5) GA10 alone without supF; (lane 6) GA10 plus supF; (lane 7) GA10 plus supF with excess unlabelled GA10; (lane 8) GA10 plus supF with excess AG10.

The reverse sequence oligomer, GA10, failed to bind to supF or to compete with AG10 for binding. AG10 linked to 4′hydroxymethyl-4,5′,8-trimethylpsoralen via a 2 carbon linker arm (pso-AG10) formed a covalent bond to labeled duplex supF DNA following UVA irradiation, whereas the reverse oligomer (pso-GA10) did not.

Targeted mutagenesis was achieved by incubating pso-AG10 with lambda supF DNA in vitro to form triplex at positions 167 to 176 of the supF gene and bring the tethered psoralen into proximity with the targeted base pair at position 167 as shown in Table 1. The numbers in the table represent the frequency of mutations seen in either the supF gene or the cI gene in the lambda supF genome following the indicated treatment. The lambda DNA at 3 nM was incubated with or without a 1000-fold molar excess of the indicated oligonucleotides (3 μM).

UVA (365 nm) irradiation of selected samples was performed at a dose of 1.8 J/cm^2. A radiometer was used to measure lamp output (typical UVA irradiance of 5-7 mW/cm^2 at 320-400 nm). The DNA was packaged in vitro into phage particles, using the method of

Hohn, *Methods in Enzymology* 68:299-309 (1979), and the phage particles wer adsorbed to E. coli and grown as individual plaques to allow genetic analyses of the supF and cI genes. AG10 bound specifically to the supF gene, whereas the reverse sequence GA10 did not bind.

CT8 (row 3), complementary to the 3′ eight nucleotides of AG10, was preincubated with psoralen-AG10 for 30 min at a 1:1 ratio to form duplex DNA and partially inhibit the ability to psoralen-AG10 to form triplex at the targeted site in the supF gene.

Photoactivation of the psoralen generated a DNA adduct, and in vitro packaging of the psoralen-AG10-lambda supF DNA complex allowed growth of the phage in bacteria to fix the adduct into a mutation. The phage particles were grown as individual plaques on a bacterial lawn to detect targeted mutagenesis in the supF gene and to measure the extent of non-targeted mutagenesis by screening for the function of an unrelated gene, the lambda repressor (cI) gene. Mutations in these genes yield colorless plaques among blue ones and clear plaques among turbid ones, respectively.

Pso-AG10 plus UVA treatment of the lambda DNA resulted in a mutation frequency of 0.233% in supF but approximately 100-fold less, 0.0024%, in cI. The specificity of the targeted mutagenesis is most likely even greater than this 100-fold difference, perhaps as much as 500-fold, considering that cI (765 bp) is a bigger target for mutagenesis than supF (184 bp) and the percentage of base pairs in the two genes at which mutations are detectable was similar. This difference in target size was demonstrated by the 5-fold difference in supF versus cI mutants induced by the reverse oligomer, pso-GA10. In addition, the reverse oligomer gave a

582-fold lower frequency of supF mutations (0.0004%) than did pso-AG10, but yielded a similar frequency of cI mutations. In fact, mutagenesis by the reverse oligomer was barely above background (untreated lambda DNA).

To partially inhibit formation of the triplex, an 8 base oligomer (CT8) complementary to 8 of the 10 bases of AG10 (5′ CCCCCTTC 3′) was preincubated at a 1:1 ratio with pso-AG10 to form a double-stranded complex. When this pre-formed complex was incubated with lambda supF and irradiated with UVA, it yielded only 0.016% supF mutations, 15-fold less than with psoralen-AG10 alone. No significant mutagenesis was produced by UVA alone (1.8 J/cm^2) in the absence of the pso-AG10 or by pso-AG10 without UVA, demonstrating the importance of activation of the psoralen by UVA and showing that triplex formation, by itself, was not mutagenic. This data provided genetic evidence for the targeted mutagenesis of the supF gene by pso-AG10.

Table 1: Targeted mutagenesis of the supF gene in lambda supF DNA produced by a psoralen-linked triplex-forming oligonucleotide (pso-AG10) plus UVA irradiation.

Treatment of lambda DNA	Sequence of oligonucleotide(s)	supF mutations per 1000 phage	cI mutations per 1000 phage
pso-AG10	pso-$^{5'}$AGGAAGGGGG$^{3'}$	2.33 (263/112.872)	0.024 (28/1,162,000)
pso-GA10	pso-$^{5'}$GGGGGAAGGA$^{3'}$	0.004 (2/504.198)	0.019 (9/483,475)
pso-AG10 plus CT8	pso-$^{5'}$AGGAAGGGGG$^{3'}$ $^{3'}$CTTCCCCC$^{5'}$	0.16 (12/72,625)	0.014 (8/557,136)
UVA alone	n.a.	<0.018 (0/55,000)	n.t.
pso-AG10 alone no UVA	pso-$^{5'}$AGGAAGGGGG$^{3'}$	<0.014 (0/69.000)	n.t.
None	n.a.	<0.003 (0/328,500)	0.009 (10/1,150,000)

Example 2: Sequence Analysis of Mutants Obtained by Targeted Mutagenesis Using the Triplex-Forming Oligonucleotide

To obtain direct evidenc for targeted mutagenesis, a series of independent mutants produced in the supF gene of the lambda vector by pso-AG10 and UVA were sequenced. The sequences of 25 such mutants are presented in Fig. 2a. All except one of the 25 mutations produced by pso-AG10 is at or near the targeted T:A base pair at position 167. 56% of the mutations consist of the same T:A to A:T transversion precisely at the targeted base pair (#167), demonstrating the specificity and reproducibility of the targeting by pso-AG10. The A:T base pair at 167 forms a triplet with the 5′ adenine to which the psoralen is tethered in AG10, and so it is the closest base pair to the psoralen. The overwhelming predominance of the T:A to A:T transversion at this site is consistent with the mutagenic action of psoralen, which tends to form adducts at pyrimidines, and especially at thymidines. It should be noted that these mutations are independent and none of the mutations represent siblings because each packaged lambda particle gives rise to a single, separate lambda plaque on the bacterial lawn.

Mutations were found to be induced in the supF gene by free 8-methoxypsoralen and UVA in other experimental systems employing shuttle vectors, as described by Glazer et al., *Proc. Natl. Acad. Sci. USA* 83:1041 (1986) and Bredberg and Nachmansson, *Carcinogenesis* 8:1923 (1987), which are incorporated by reference herein. This compiled data demonstrates that free psoralen can form adducts and induce mutations at many different sites in supF apart from base pair 167. The scattered distribution of mutations is in contrast

with the specific mutagenesis induced by the triplex-forming pso-AG10. Although sev ral of the mutations list d in Fig. 2b fall in the region of the homopurine/homopyrimidine run at positions 167 to 176, none of them occur at position 167. Neither of the two mutations induced by the reverse oligomer, pso-GA10, were found to occur at base pair 167.

The spectrum of the mutations produced by pso-AG10 indicates that almost all were targeted by the triplex-forming oligonucleotide. Although a majority of the mutations were at the targeted position 167 and consisted of the same T:A to A:T transversion, several mutations were at base pairs nearby to position 167. It is possible that the psoralen moiety, tethered to AG10 on a 2 carbon linker arm, may occasionally reach beyond the T:A base pair at 167 to form adducts at nearby pyrimidines, giving rise to mutations. It is also possible that even if an adduct is formed at position 167, the bacterial polymerase and repair enzymes that fix the adduct into a mutation may generate mutations at nearby sites during repair and replication while at the same time repairing or bypassing the adduct at 167. The occurrence of several mutations that involve base changes at two adjacent base pairs (166 and 167 in all 3 instances) supports the notion that an adduct at position 167 can cause a change at a nearby position. The rare non-specific mutagenesis by pso-AG10 (and the very small amount of mutagenesis by pso-GA10 that is above background) may result from the potential ability of the psoralen molecule, in spite of being tethered to the oligonucleotide, to intercalate into and form adducts at random sites in the DNA. A reduction of this non-specific activity may be achieved by

reducing the reach and the degrees of freedom of the psoralen by attaching it to th triplex-forming oligonucleotide by a shorter tether, such as a n carbon linker arm, or by direct linkage of the psoralen to the nucleotide in the triplex-forming oligonucleotide by direct photoactivation of free psoralen to bind to the oligonucleotide, and the purification of the desired product.

This experiment achieved a targeted mutation frequency of 0.233%.

Example 3: Covalent Linkage of Psoralen to an Oligonucleotide.

The mutagen, 5-aminomethyl-8-methoxypsoralen, was covalently linked to an oligonucleotide.

5-aminomethyl-8-methoxypsoralen (5am8mop, HRI Associates, Emeryville, CA) was mixed with the linker, sulfo-m-maleimidobenzoyl-N-hydroxysuccinimide ester (sulfo-MBS, Pierce Chemical Co., Rockville, IL) in 0.05 M phosphate buffer, pH 8, with a 5am8mop to sulfo-MBS molar ratio of 1:40. The mixture was stirred at room temperature for 30 minutes while protected from light in accordance with the methods of Liu et al., *Biochem.* 18:690-697 (1979) and Kitagawa and Ailawa, *J. Biochem.* 79:233-236 (1976), and the instructions of the Pierce Immunotechnology Catalog and Handbook, 1992-93 edition, pages A16-A17. The initial run was made using 1 ml total volume and 1 mM 5am8mop, with the reaction scaled up and optimized as needed.

The modified 5am8mop was purified by HPLC using a modification of standard conditions used in the analysis of 8-methoxypsoralen as described by Gasparro et al., *J. Invest. Derm.* 90:234-236 (1988). The initial conditions were: a Regis Rexchrom™ phenyl 15 cm HPLC column running a gradient between acetonitrile and either water or

0.05 M, pH 4.5 ammonium acetat buff r. A linear gradient was run from 10% acetonitrile to 60% acetonitrile over 50 minutes. When buffer was needed in the initial purification run, the sample was collected off the HPLC, evaporated, and desalted by passing it through the HPLC again with an acetonitrile:water gradient mixture. The detector was a SpectraFocus™ scanning UV detector with wavelengths from 220 to 360 sampled. The detector was connected to a Pharmacia Frac-100™ fraction collector.

The purified, modified 5am8mop was then reacted with an oligonucleotide containing an -SH tether by mixing equimolar amounts of modified 5am8mop with the oligonucleotide in 0.05 M phosphate buffer, pH 7-7.5 at room temperature for three hours while protected from light.

The oligonucleotide tethered to 5am8mop was then purified by HPLC using a modification of the method of Gasparro et al., *Antisense Res. Dev.* 1:117-140 (1991). A Nest Group MRPH 10 cm HPLC column running a linear gradient of 5% to 20% acetonitrile over 40 minutes between acetonitrile and 0.2 M pH 5.9 triethylammonium acetate buffer was used.

Example 4: Targeted Mutagenesis of SV40 DNA Using Triple Helix-Forming Oligonucleotides.

The following was performed to investigate targeted mutagenesis of SV40 DNA transfected into monkey cells. In these experiments, the site-specific triplex formation was designed to deliver the psoralen to the targeted site in the SV40 DNA, UVA irradiation was used to activate the psoralen to form adducts at that site, and repair and replication of the viral genomes in the monkey cells fix the adducts into mutations. These results demonstrate that targeted mutagenesis occurs even more efficiently in mammalian cells (6%

of SV40 genomes incurred targeted mutations) than in bacteria (0.2%).

MATERIALS AND METHODS

Oligonucleotides and vectors. Psoralen-linked oligonucleotides were obtained from either Oligos Etc. (Wilsonville, OR) or M. Talmor (Yale University, New Haven, CT) with materials from Glen Research (Sterling, VA). The psoralen is incorporated in the oligonucleotide synthesis as a psoralen phosphoramidite, resulting in an oligonucleotide linked at its 5′ end via a two carbon linker arm to 4′-hydroxymethyl-4,5′,8-trimethylpsoralen, as illustrated in Fig. 3. The sequences of oligonucleotides used in this study include AG10 (5′AGGAAGGGGG3′) and GA10 (5′GGGGGAAGGA3′). SV40 shuttle vector pSP189 was constructed by and obtained from Dr. Michael Seidman (Otsuka Pharmaceuticals, Bethesda, MD).

Triplex binding assays. Binding assays were carried out for 2 hours at 37°C in 10% sucrose, 20 mM $MgCl_2$, 10 mM Tris (pH 8.0), and 1 mM spermidine in a 10 μl volume. The 250 bp supF target was generated from lambda supF using the polymerase chain reaction.

Protection assay using PCR amplified supF target. The 250 bp supF target (70 nM) was incubated with a 100-fold molar excess of psoralen-AG10 as described for the binding assay. Irradiation of samples was performed at a dose of 1.8 J/cm^2. A radiometer was used to measure the lamp output (typical UVA irradiance of 5-7 mW/cm^2 at 320-400 nm). Following the binding and irradiation steps, samples were digested for 2 hours at 37°C with *Hinf* I. Loading buffer was added and samples were heated 10 minutes at 55°C, and run for 1 hour on a 4.5% Nusieve gel in TAE buffer at 80 v (10 v/cm). An analysis by agarose gel electroph resis f *Hinf*

I digestions of the 250 bp supF gene PCR fragment under various conditions was performed. A faint band corresponding to a size of 150 kDa was present in Lane 1, which contained no psoralen-AG10 and no UVA; a band corresponding to a size of 150 kDa was present in lane 2, containing UVA alone (no psoralen-AG10); a band corresponding to a size of 150 kDa was present in lane 3, psoralen-AG10 and UVA; a band corresponding to a size of 150 kDa was present in lane 4, psoralen-AG10 alone (no UVA); a band corresponding to a size of approximately 300 kDa was present in lane 5, undigested supF PCR fragment; lane 6 contained the size markers (100 bp ladder) and bands were present at 100, 200 and 300 kDa.

Protection assay using SV40 vector DNA target. The binding and irradiation were carried out as described above, except that pSP189 was used as the supF target at a concentration of 50 nM and psoralen-AG10 was added at ratios of oligomer to vector of from 1:1 to 1000:1. Irradiation and gel conditions were as described above.

Colony hybridization. Ampicillin resistant colonies of SY204 carrying shuttle vector plasmids with supF gene mutations, along with appropriate control colonies, were grown on LB/ampicillin plates and transferred onto replica nylon filters for additional growth and *in situ* lysis to allow colony hybridization by standard methods. The DNA was fixed to the filters by UV crosslinking, and the filters were incubated in 6X SSC, 5X Denhardt's solution, 0.5% SDS, and 5 x 10^5 cpm/ml of ^{32}P-labeled oligonucleotides at 42°C for 18 hours. The filters were washed in 1X SSC and 0.1% SDS for 30 minutes at 25°C and then in 1X SSC and 0.1% SDS at 42°C for 2 hours. These conditions were empirically determined to all w discrimination

between binding of the wild type probe (5′ GGT TCG AAT CCT TCC CCC 3′) and the 167 mutant probe (5′ GGT TCG AAA CCT TCC CCC 3′). Binding of the oligonucleotide probes was determined by autoradiography.

SV40 mutagenesis. The SV40 vector DNA (pSP189) at 80 nM was incubated with psoralen-AG10 or psoralen-GA10 (ranging from 2 to 1000-fold molar excess) and irradiated as described above. The oligonucleotide-plasmid complex was then transfected into monkey COS-7 cells (ATCC #1651-CRL) using cationic liposomes (DOTAP, Boehringer Mannheim, Indianapolis, IN) at a final concentration of 5 μg/ml in the culture dish. The DNA/oligo/liposome mixture was added dropwise to the cell culture dish with swirling. The following day, the media containing the liposome mixture was replaced by fresh media. Following 48 hours to allow repair and replication, SV40 vector DNA was harvested from the COS cells by the Hirt lysate procedure. Genetic analysis of the supF genes in the SV40 vector was carried out by transformation of E. coli SY204 [*lacZ*125(Am)] to ampicillin resistance by electroporation using 12-150 ng of *Dpn* I digested Hirt lysate DNA and a Bio-Rad Gene Pulser apparatus equipped with a Pulse Controller (Bio-Rad, Richmond, CA). Mutants were identified by growth in the presence of 65 μg/ml IPTG and 80 μg/ml X-Gal, as described by (Glazer et al., *Mol. Cell Biol.* 7:218-224 (1987)). These transformants were counted and the mutants (white colonies) were streaked for single colonies.

DNA sequencing. DNA was prepared for sequencing by isolating DNA from a 3 ml bacterial culture using a Promega Magic Miniprep kit (Promega, Madison, WI). DNA sequence data was obtained by direct chain

termination sequencing of the plasmid DNA using automated methods.

RESULTS

Strategy for targeted mutagenesis in SV40. An SV40-based shuttle vector (pS189) was used to assay for targeted mutagenesis. This vector contains both the SV40 and the pBR328 origins of replication, plus the β-lactamase gene for ampicillin resistance, to allow episomal replication in both mammalian cells and bacteria (Fig. 3). It also carries the supF gene, an amber suppressor tyrosine tRNA gene of E. coli, as a marker gene for mutagenesis studies.

In this vector system, the SV40 DNA, after appropriate treatment, is introduced into monkey COS cells where repair and replication can occur, producing mutations indicative of mammalian processing of DNA damage. The small, circular vector DNA is recovered from the cells by biochemical separation from the chromosomal DNA (Hirt lysate, Hirt et al., *J. Mol. Biol.* 26:365-369 (1967)), and it is used to transform E. coli carrying the lacZ (amber) mutation to allow analysis of supF gene function by scoring colonies for β-galactosidase activity (produced via suppression of the amber mutation in lacZ) in the presence of the chromogenic substrate, X-gal. Vectors with wild type supF genes yield blue colonies; those with mutations in supF produce white ones. In order to eliminate misleading data that might arise from viral DNA that was not replicated or repaired in the mammalian cells, the viral DNA is digested before bacterial transformation with the enzyme *Dpn* I which will restrict DNA that has not been methylated by the mammalian pattern at its recognition site.

The design of the initial experiments to target mutations to SV40 DNA is illustrated in Fig. 3. A 10 base pair region of the supF gene (bp 167-176) was identified as a site amenable to triplex formation because of the homopurine/homopyrimidine run there. Since this run was G-rich, the purine motif for triplex formation was selected (Beal and Dervan, *Science* 251:1360-1363 (1991)), and an oligonucleotide, 5'AGGAAGGGGG3' (AG10) was synthesized based on this motif. A psoralen derivative, 4'-hydroxymethyl-4,5',8-trimethylpsoralen, was attached to the oligonucleotide by a phosphodiester linkage at the 5' adenine via a two carbon linker arm, with the goal of directing mutations to base pair 167. This is the base pair with which that 5' adenine binds in the predicted triple helix. Note that the psoralen-AG10 oligonucleotide is oriented anti-parallel to the purine-rich strand in the duplex DNA. To achieve targeted mutagenesis, the pSP189 DNA is incubated with the psoralen-linked oligonucleotide (psoralen-AG10), treated with long wave ultraviolet light (UVA) to activate the psoralen to form a pre-mutagenic adduct on the thymidine in base pair 167, and then transfected into COS-7 cells. After a 48 hour period to allow repair and replication, the viral DNA is isolated from the monkey cells, subjected to digestion with *Dpn* I, and used to transform E. coli. The frequency of supF mutations is determined, and representative samples of supF mutant clones are collected for further analysis.

Site-specific formation of triplex DNA. This experiment demonstrate the ability of psoralen-AG10 to bind specifically to the intended site within the supF gene using a restriction enzyme protection assay. In this assay, psoralen-AG10 was found to

bind site-specifically to duplex supF DNA following UVA irradiation, blocking restriction enzyme digestion at the one *Hinf* I site (bp 164-168) that overlaps the triplex target site (167-176) but not at the other *Hinf* I site in supF (bp 129-133). This is diagrammed in Fig. 4 and was demonstrated by the electrophoretic gel (not shown) which showed site-specific formation of triplex DNA in the SV40 vector as a function of the ratio of oligonucleotide to SV40 DNA. The SV40 vector containing the supF target gene (50 nM) was incubated with psoralen-AG10 at ratios of oligomer to vector of from 1:1 to 1000:1, irradiated with 1.8 J/cm^2 of UVA, digested with *Hinf* I, and run on a 4.5% Nusieve gel. Lane 1, undigested plasmid DNA; lane 2, no psoralen-AG10 prior to digestion; lanes 3-7, increasing ratios of psoralen-AG10/SV40 DNA as indicated above each lane; lane 8, 100 bp size markers (BRL-Gibco). Because the sequences flanking the supF gene in the SV40 DNA differ from those in the PCR fragment presented in the gel described above and since there are multiple *Hinf* I sites in SV40, the pattern of bands is more complex than those in the gel described above. However, this gel indicates the position of the fragment resulting from shielding of the *Hinf* I site at bp 164-168 by triplex formation with a band between size markers 100 and 200.

Digestion of the unprotected 250 bp supF PCR fragment with *Hinf* I yields three fragments of sizes 150, 65, and 35 (lane 1), in contrast with the uncut fragment of 250 bp (lane 6). Incubation of the supF fragment with psoralen-AG10 along with photoactivation with UVA (lane 3) results in protection of the *Hinf* I site at bp 164-168 but not the one at bp 129-133, as demonstrated by the appearance of the 100 bp fragment instead of the 65

bp and 35 bp fragments. UVA-induced covalent adduct formation is required for restriction enzym protection, since psoralen-AG10 alone is not sufficient to prevent *Hinf* I digestion (lane 4). In the absence of psoralen-AG10, UVA light had no effect on *Hinf* I digestion (lane 2). In similar experiments, no protection from *Hinf* I cutting was seen using psoralen-GA10, the reverse sequence oligomer linked to psoralen. This data demonstrates site-specific formation of triplex DNA by psoralen-AG10, with covalent modification of the supF gene fragment occurring at the targeted site following UVA irradiation of the psoralen-AG10/supF complex.

Similar experiments were performed to assay for site-specific binding of psoralen-AG10 to bp 167-176 in the supF gene within the SV40 vector itself. In these experiments, varying ratios of oligonucleotide to vector DNA were employed in order to examine basic parameters of the triplex binding to the viral genome. The above described gel indicates that *Hinf* I protection at the targeted site is almost complete at a 10:1 ratio of oligonucleotide to vector, as judged by the appearance in the ethidium bromide stained agarose gel of a band at 125 bp (arrow) and the disappearance of the band at 90 bp. Ratios of 100:1 and 1000:1 similarly yielded near complete protection, whereas the lower ratios of 1:1 and 2:1 gave only partial protection. These results are consistent with the mutagenesis experiments, described below.

Targeted mutagenesis of SV40 vector DNA passaged in COS cells. Experiments to induce targeted mutagenesis in SV40-vector DNA using triplex-forming oligonucleotides were carried out as shown in Fig. 3. Psoralen-linked oligonucleotides were

incubated with SV40 vector DNA, exposed to 1.8 J/m^2 UVA light, and transfected into COS cells. After two days to allow repair and replication to occur, the vector DNA was rescued from the cells and used to transform bacteria to facilitate genetic analysis of the supF gene. The effect of psoralen-AG10, which binds site-specifically to the supF gene in the vector, in inducing supF mutations was compared to that of psoralen-GA10, which shows no specific binding. Various ratios of oligonucleotide to vector DNA were used in order to investigate parameters that might affect the specificity and the efficiency of the process of targeted mutagenesis in the monkey cells. Table 2 presents the data from these experiments. Targeted mutations in the supF gene were produced in the SV40 genome at a frequency as high as 7.3 % using psoralen-AG10 at a molar ratio of oligonucleotide to vector DNA of 1000 to 1. At this same ratio, psoralen-GA10 produced a small amount of mutagenesis above background (0.5% versus 0.07%). At the lower ratios tested, however, the reverse oligomer yielded no significant mutagenesis above the background frequency in the assay, whereas, at these lower ratios, psoralen-AG10 still generated a high frequency of mutations in supF (as high as 6.4% for the 10:1 ratio versus 0.06% for psoralen-GA10 at 10:1 and 0.07% for untreated vector DNA). This demonstrates mutagenesis specifically targeted to the supF gene in the SV40 vector by psoralen-AG10 but not by psoralen-GA10. This frequency of targeted mutagenesis in SV40, in the range of 6% to 7%, is 30-fold higher than that seen in previous experiments to target the supF gene in bacteriophage lambda grown in E. coli (0.23 %, Havre et al., *Proc. Natl. Acad. Sci.* 90:7879-7883 (1993)), and it suggests that the monkey cells more

efficiently process the pre-mutagenic lesion of the psoralen/oligonucleotide adduct into a mutation, via either error-prone repair or bypass replication.

In control experiments (Table 2), UVA irradiation of the SV40 DNA, in the absence of the psoralen-linked oligonucleotides, produced no mutagenesis above background. Similarly, the treatment of the SV40 DNA with the oligomers but without UVA irradiation was not mutagenic.

Sequence analysis of targeted mutations. A set of 20 mutants generated in the supF gene in the SV40 vector by psoralen-AG10 (at the 1000:1 ratio) and UVA light were subjected to DNA sequence analysis. The results of this analysis are shown in Fig. 5A. Of the 20 mutations analyzed, 11 consist of the same T:A to A:T transversion at base pair 167 occurring over and over again. This is the precise base pair to which the mutations were targeted by psoralen-AG10, as diagrammed in Figs. 5A and 5B. The finding that 55% of the sequenced mutations consisted of the exact same base change at the targeted base pair suggests that the intended base change (T:A to A:T at bp 167) was produced in over 4% of all the viral genomes. The other mutations analyzed included 3 point mutations at base pairs adjacent to the targeted base pair and 6 small deletions including or abutting that base pair. These likely arise from variations in the processing, repair, or replicative by-pass of the triplex-directed lesion at bp 167 as the SV40 DNA is replicated in the monkey cells. It is also possible that the psoralen molecule, tethered to the oligonucleotide by a 2 carbon linker arm, has sufficient reach and degrees of freedom to form adducts at nearby base pairs. Improved mutational specificity may be achieved by reducing the length

of the linker arm. In Fig 5B, the published sequences of supF mutation produced in this same vector system using free 8-methoxypsoralen are presented for comparison. Not only are these mutations more scattered, but also none were found to occur at base pair 167.

In the analysis of mutagenesis in SV40 vectors, it is often difficult to determine if identical mutations arose independently or if they were the result of a single mutational event amplified by subsequent vector replication. In order to exclude the possibility that such siblings mutations were isolated in these experiments, use was made of an advantageous feature of the pSP189 system, in which over 100,000 different, random-sequence 8 base pair oligonucleotides were cloned into a region of pSP189 next to the supF gene. The vector DNA is prepared *en masse* from the library of vector clones containing the different 8 bp sequences. In this way, at the same time that the sequence of the supF gene in a mutant vector is ascertained, the 8 base pair signature sequence in that particular plasmid molecule can also be identified by reading a few extra bases further into the sequence data. This enables comparisons between the 8 bp signature sequences in plasmids bearing the same supF mutation to see if they are siblings from the same mutational event or if they are independent mutations. Based on this analysis, it was determined that all 20 of the mutations presented here arose independently.

In order to strengthen and confirm these results, a larger sample of supF mutations produced in the SV40 vector by psoralen-AG10 and UVA light was analyzed by an alternate method based on the expected high proportion of T:A to A:T transversions at bp 167. Instead of direct

sequencing, a technique o. ifferential oligonucleotide hybridizat ons was used (Sidransky et al., *Science* 252:706-7((1991)). In this assay, undertaken in an ef rt to streamline mutant analysis, ampicillin resis nt bacterial colonies containing mutant supF ger were grown on nylon filters to allow nucleic a d hybridizations. Duplicate filters were inc bated with ^{32}P-labelled, 18 base oligonucleotides t at either matched the wild type sequence or mat ed the position 167 T:A to A:T mutant sequence. T hybridizations were carried out by standard me ods under conditions empirically determined to stringent enough to allow differentiation bet n mutant and wild type sequences. Of the 19 col ies assayed in this particular experiment, 9 s owed hybridization specific to the mutant pr e. None showed hybridization to the wild pe probe, except for the positive control in th upper right corner. For the 9 colonies that bi to the 167 probe, this supports the validity of assay. For the other 10 that did not bind to t mutant probe either, the lack of binding to t wild type probe suggests that they either ave different mutations at bp 167 (not T:A to A:T) or have mutations near bp 167, within the 18 bp r gion covered by the probes, causing mismatche ith both the wild type and mutant oligonucleotid . A total of 42 mutants generated by psoralen-AG1 were nalyzed by this method (including the 20 bject to sequence analysis), and 22 (52%) w found to carry the T:A to A:T mutation at bp 167. All of the rest were judged to have different ations at or near the targeted base pair, becau neither the mutant nor wild type probe hybridize to them. The validity of this assay was support by the 100% agreement with the sequencing data. These results extend the

direct sequencing data ar demonstrate further the targeted mutagenesis of S vector DNA. Taken together, the data sugges that almost all of the mutations produced by psc len-A310 are at or within a few bases of the targeted base pair, and at least 50% consist of t same T:A to A:T transversion at that site These results demonstrate efficient pro of specific, reproducible, and predict mutations at a targeted base pair in SV4 DNA passaged in monkey cells.

TABLE 2. Targeted mutagenesis in SV40 DNA

Treatment of SV40 vector DNA[a]	Ratio of oligo to vector	% mutants[b]	Mutants per total colonies
None	n.a.	0.07	6 / 8,190
psoralen-AG10[c] no UVA	1000:1	≤0.06	0 / 1700
psoralen-GA10[c] no UVA	1000:1	≤0.07	0 /1500
UVA alone	n.a.	0.06	5 / 8,427
psoralen-AG10[c]	2:1	2.5	148 / 5,869
	5:1	4.3	118/ 2,734
	10:1	6.4	381 / 5,995
	1000:1	7.3	633 / 8,643
psoralen-GA10[c]	2:1	0.07	3 / 4,397
	5:1	0.3	11 / 8,230
	10:1	0.06	4 / 6,800
	1000:1	0.63	92 / 14,670

[a] Except where indicated, all samples received 1.8 J/cm^2 of UVA irradiation.

[b] The values represent the frequency of mutations seen in the supF gene within the pSP189 SV40 vector.

[c] Psoralen-AG10 forms a site-specific triple strand at bp 167-176 of the supF gene within pSP189; the reverse sequence oligomer, psoralen-GA10, does not.

Example 5: Intracellular Targeted Mutagenesis.

An experiment demonstrating targeted mutagenesis of an SV40 vector mediated by intracellular triple helix formation in monkey CS cells was performed. The results demonstrated that specific, reproducible mutations can be produced in viral genomes replicating in monkey cells by treatment of

the cells with psoralen-linked oligonucleotides followed by photoactivation of the psoralen with long wavelength UV light (UVA). Using a set of modified target sites in the *supF* gene and a series of oligonucleotides, this experiment demonstrated that targeted mutagenesis in an *in vivo* assay depends on the specificity of the mutagenic oligonucleotide for the target site and on the strength of the oligonucleotide binding. The results also provide a preliminary analysis of both the time dependence and the concentration dependence of the oligonucleotide-directed targeting within mammalian cells.

Materials and Methods.

Oligonucleotides and vectors. Psoralen-linked oligonucleotides were obtained from Oligos Etc. (Wilsonville, OR). The psoralen was incorporated into the oligonucleotide synthesis as a psoralen phosphoramidite, resulting in an oligonucleotide linked at its 5′ end via a two-carbon linker arm to 4′-hydroxymethyl-4′,5′.8-trimethylpsoralen, as illustrated in Fig. 6. The sequences of psoralen-conjugated oligonucleotides used in this experiment include:

AG10 5′ AGGAAGGGGG 3′

AGT20 5′ AGGAAGGGGGGGGTGGTGG 3′

AGT30 5′ AGGAAGGGGGGGGTGGTGGGGAGGGGGAG 3′

AGT43 5′ AGGAAGGGGGGGGTGGTGGGGAGGGGG
AGGGGGAGGGGGAG 3′

SV40 shuttle vectors, pSupFG1a and pSupFG2 were derivatives of pSP189 (described above) and carried new triplex-binding sites which were engineered into the *supF* gene. The modified *supF* genes were constructed by inserting synthetic oligonucleotides into the *Xho*I to *Eag*I sites in the original *supF* gene using standard techniques as described by Sambrook et al., MOLECULAR CLONING: A LABORATORY MANUAL,

second edition, Cold Spring Harbor Laboratory Press, New York (1990) (which is incorporated by reference herein).

Triplex binding assays. Two complementary 57-mers which contain the sequence corresponding to bp 157 to 213 of *supFG1a* were synthesized. Both oligomers were labeled with y-[^{32}P]-ATP. The duplex DNA was prepared by mixing both 57-mers at a ratio of 1:1 in TE buffer and incubating at 37°C for two hours. A fixed concentration of duplex DNA (1 x 10^{-10} M) was incubated with increasing concentrations of the psoralen-linked oligomers in 10 μl of 10 mM Tris (pH 7.4), 1 mM spermidine, and 20 mM $MgCl_2$ at 37°C for two hours. UVA irradiation (1.8 J/cm^2 of broad band UV light centered at 365 nm, irradiance of 5 mW/cm^2) was used to generate photoadducts and thereby covalently link the oligomers to their targets. The samples were mixed with 90 μl of formamide, and 20 μl of each sample was analyzed on an 8% polyacrylamide denaturing gel containing SDS and 7 M urea. Phosphor-imager (Molecular Dynamics, Sunnyvale CA) was used for quantitation of the reaction products. The concentration at which triplex formation (as indicated by the generation of specific photoadducts) was half-maximal was taken as the equilibrium dissociation constant (Kd).

Mutagenesis protocol. Transfection of the cells with plasmid and treatment of the cells with the psoralen-oligonucleotide was performed as follows. Monkey COS-7 cells were obtained from the ATCC (1651-CRL) (Rockville, MD). The COS cells at 70% confluence were washed with PBS-EDTA, treated with trypsin and incubated at 37°C for five minutes. The cells were resuspended in DMEM/10%FCS and were washed three times by centrifugation at 900 rpm for five minutes (4°C) using a Sorvall

RT6000D centrifuge. The cells were resuspended at 1 x 10^7 cells/ml. The plasmid DNAs were added at 3 µg DNA/10^6 cells and the cell/DNA mixtures were left on ice for 10 minutes. Transfection of the cells was performed by electroporation using a Bio-Rad gene pulser at a setting of .5 µF/250 W/250 V in the 0.4 cm cuvette. Following electroporation, the cells were kept on ice for 10 minutes. The cells were diluted with growth medium, washed, and transferred to 37°C for 30 minutes. At this point, the cells were either further diluted and exposed to the oligonucleotides in growth medium or washed, diluted further, and allowed to attach to dishes for twelve hours, washed again with PBS/EDTA, trypsinized, washed three times with growth medium, and finally exposed to the oligonucleotides in suspension. In each case, the psoralen-conjugated oligonucleotides were added to cells in suspension, which were then incubated at 37°C with gentle agitation every fifteen minutes. UVA irradiation was given at a dose of 1.8 J/cm^2 at the indicated times. The cells were further diluted in growth medium and allowed to attach to plastic dishes at a density of 1 x 10^6 cells per 15 cm^2 dish.

After 48 hours, the cells were harvested for vector DNA isolation using a modified alkaline lysis procedure. The cells were resuspended in 100 µl of cell resuspension solution (50 mM Tris/HCl, 10 mM EDTA, pH 8.0; 100 µg/ml RNase A) and 100 µl of cell lysis solution (0.2 M NaOH, 1% SDS) was added. After five minutes at room temperature, 100 µl of neutralization solution (3 M potassium acetate, pH 5.5) was added. A fifteen minute room temperature incubation was followed by centrifugation in a microcentrifuge for ten minutes. The clear supernatant was extracted with an equal volume of phenol/chloroform (1:1) once,

and the DNA was precipitated with 2.5 volumes of ethanol at -70°C for ten minutes. The DNA was collected by centrifugation for ten minutes, washed with 70% ethanol once, and allowed to air dry for five minutes at room temperature. The DNA was digested with *Dpn* I and RNase A at 37°C for two hours, extracted with phenol/chloroform, and precipitated with ethanol. The DNA pellet was dissolved in 10 µl of TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0) and 1 µl of viral DNA was used to transform *E. coli* SY204 or DH10B by electroporation. The transformed *E. coli* cells were plated onto LB plates containing 50 µg/ml of ampicillin, 100 µg/ml of X-gal and 1 M of IPTG and were incubated at 37°C overnight. The mutant colonies and the total colonies were counted. The mutant colonies were purified and the plasmids were isolated for DNA sequence analysis.

DNA sequencing. The single colonies of purified mutants were picked into 5 ml of L broth containing ampicillin (50 µg/ml) and were incubated at 37°C for 16-20 hours by shaking at 250 rpm. Cells from 3 milliliters of culture were collected by centrifugation. Isolation of plasmid DNA was accomplished using the Wizard plasmid miniprep DNA purification system (Promega, Madison, WI). 1.5 µg of plasmid DNA was used for DNA sequencing using an ABI cycle-sequencing kit in accordance with the manufacturer's instructions (Applied Biosystems Inc., Foster City, CA) using standard methods. The sequencing primer was chosen to bind to the ß-lactamase gene just upstream of the *supF* gene in the vector.

Results.

Mutagenesis assay. The design of the assay system used to study targeted mutagenesis of an SV40 vector *in vivo* in Monkey cells is shown in

Fig. 6. In addition to the SV40 sequences necessary to enable replication in mammalian cells, the SV40 shuttle vector contains the *supF* gene, a suppressor tRNA gene of *E. coli*, as the target gene for mutagenesis. It also contains a portion of the pBR327 replication origin and β-lactamase gene for growth and selection in bacteria. Following introduction of the vector DNA into the COS cells by electroporation, the cells were incubated in the presence of the psoralen-linked oligonucleotides, which are designed to bind to *supF* gene sequences such that the psoralen is delivered to the intended intercalation site at bp 166 1. After allowing a variable time for entry of the oligomers into the cells and for triplex formation, the cells were irradiated with UVA to activate the psoralen to form photoadducts and consequent mutations at the targeted site in the *supF* gene. Another 48 hours were allowed for repair and/or replication. The vector DNA was harvested from the cells by an alkaline lysis procedure and used to transform *lacZ* (amber) *E. coli* to detect mutations in the *supF* gene that occurred in the COS cells. Prior to transformation, *Dpn I* digestion of the vector DNA was used to eliminate unreplicated vector molecules lacking the mammalian methylation pattern and thereby preclude misleading results resulting from unprocessed input molecules.

Insufficient *in vivo* targeting by a 10-mer. In the experiments described above regarding *in vitro* experiments, a 10 base pair region of the *supF* gene, bp 167-176, was targeted for triplex formation in the anti-parallel motif using a 10 nucleotide psoralen-linked oligomer, Iso-AG10. In these experiments, the triplex formation was allowed to occur during a two hour incubation *in vitro* in an optimized Mg^{2+} containing buffer,

followed by *in vitro* UVA irradiation and then transfection of the vector-oligonucleotide complex into the cells as described above. Mutations were generated in the target gene at a frequency of approximately 6 to 7%. In subsequent experiments, the SV40 vector and pso-AG10 were co-transfected into the cells, with UVA irradiation of the cells two hours later. This protocol yielded targeted mutations in 2.1% of the vector molecules, a frequency in the range of that observed in the above-described *in vitro* work. However, attempts to use pso-AG10 to target mutations to the *supF* gene by treatment of cells that were pre-transfected with the pSP189 vector in the protocol outlined above (Fig. 6), resulted in detection of little mutagenesis (0.05%) above background (0.02%).

Construction of novel *supF* genes. Based on this result and based on the reported binding constants for oligonucleotides of various lengths, a modified *supF* gene containing 30 base pair polypurine/polypyrimidine segments amenable to triplex formation was constructed because it was believed that the binding affinity for triple helix formation by the 10-mer (pso-AG10) might be insufficient to achieve significant interactions *in vivo*. Novel gene sequences were incorporated into the *supF* gene to test a series of potential triplex forming oligonucleotides using position 167 as the targeted psoralen intercalation site.

Assay systems were developed to allow the characterization of factors affecting *in vivo* triplex formation and to examine the possibility of intracellular mutation targeting.

The pSP189 vector contains a 93 bp segment encompassing the sequences coding for the mature tRNA between unique *Xho*I and *Eag*I sites. Using synthetic oligonucleotides, this 93 bp stretch was

replaced with novel sequences. The design of a new *supF* gene, *supFG1a*, is illustrated in Fig. 7. To eliminate one interruption at bp 167 in the polypurine/polypyrimidine run, an A:T to C:G transversion was incorporated into the synthetic fragment, along with a compensatory T:A to G:C change at bp 101 to maintain base pairing in the amino acid acceptor stem of the mature tRNA. In addition, a 13 bp polypurine/polypyrimidine sequence was inserted between bp 183 and 184 to extend the length of the polypurine/polypyrimidine run in the gene to 30 bp. Since the 3' CCA sequence at positions 181-183 comprises the 3' terminal amino acid acceptor site of the tRNA, the new sequences 3' to position 183 do not affect the mature tRNA molecule and so do not alter the phenotype of the gene. Following annealing and ligation of the synthetic oligonucleotides into the vector, constructs containing functional suppressor genes were identified by transformation of *lacZ* (amber) bacteria. The sequence of the new *supFG1a* gene was confirmed by direct DNA sequencing of the vector DNA. The new construct contains a 30 bp polypurine site with just two interruptions. By a similar method, *supFG2* was synthesized, containing a 43 bp polypurine site, also with just two interruptions, as shown in Fig. 8.

Triplex binding. To compare triplex formation at the polypurine sites in *supF*, *supFG1a*, and *supFG2*, a gel mobility shift assay was used in which synthetic DNA fragments of the 24 bp (matching bp 160-183 in the *supF* gene), 57 bp (matching bp 157-213 in *supFG1a*) or 63 bp (matching bp 159-217 in *supFG2*) were used as duplex targets for triplex formation by addition of oligonucleotides. Corresponding oligonucleotides pso-AG10, pso-AGT30, and pso-AG were designed to

bind in the anti-parallel motif to the 0, 30, and 43 bp sites in these genes [illegible] G220, designed to bind to bp 167-186 in *supF* [illegible] tested. Fixed concentrations of radiolabelled duplex target DNA were incubated with increasing concentrations of the psoralen [illegible] oligomers to assay for triplex formation over a range of concentrations. UVA irradiation (1.8 J/cm2) was used to generate photoadducts and thereby covalently link the mutagenic oligonucleotides to their targets, ensuring that subsequent manipulation of the samples will not alter the apparent binding. The samples were then analyzed by denaturing gel electrophoresis and autoradiography.

Analyses of the binding of pso-G10 and pso-AGT30 to their respective sites in the *supF* and *supFG1a* genes demonstrated that activation of the triplex molecules with a dose of UVA leads to a high proportion of cross-links of psoralen-mutagenic oligonucleotide covalently to both strands of the duplex target via a psoralen interstrand cross-link - the XL band, and to a small proportion of psoralen monoadduct (the psoralen-mutagenic oligonucleotide linked covalently linked to just one strand of the duplex via psoralen monoadduct formation - the MA band). The percentage of the sample containing the sum of the XL and MA bands is proportionate to the extent of triple helix formation. Other oligonucleotides that cannot form triplex at the target site are used in this assay. XL or MA bands are visualized (not shown) as adducts can be taken as indicative of triplex formation. As the concentration of the triplex forming oligonucleotide is increased, the proportion of the target duplex bound as either triplex, as opposed

to unbound, increases. As can be seen from the graphical analysis in Fig. 9, the concentration dependence of triplex formation by pso-AG10 and by pso-AGT30 are quite different. A useful measure of triple helix formation is the equilibrium dissociation constant, K_d, which can be estimated as the concentration of mutagenic oligonucleotide at which triplex formation is half-maximal. For pso-AG10 the K_d is 8×10^{-7} M, whereas for pso-AGT30, the K_d is 3×10^{-9} M, a 270-fold difference. The greater affinity for triplex-formation by pso-AGT30 is consistent with the results of other studies correlating oligonucleotide length and binding affinity, and it places the affinity of pso-AGT30 for binding to *supFG1a* in the range of physiologic interactions. Similar analyses revealed a K_d of 1×10^{-8} M for pso-AGT43 binding to *supFG2* and 8×10^{-9} M for pso-AGT20 binding to *supFG1a*.

Targeted mutagenesis ***in vivo***. Using pSupFG1a as a target vector, experiments were conducted to study intracellular mutation targeting by three psoralen-linked oligonucleotides, all of which were designed to form a triple helix with sequences in *supFG1a* and to deliver the tethered psoralen to bp 166-167. However, the oligomers differed in length and binding affinity for triplex formation as shown in Table 3 below. In these experiments the oligomers were added to the cells approximately one hour following electroporation with the vector DNA. The cells were then irradiated with UVA either two hours or eight hours after oligonucleotide addition. At both the two and eight hour time points, the extent of mutagenesis of *supFG1a* gene was seen to depend on the relative strength of triplex formation by the respective mutagenic oligonucleotide. At two hours pso-AG produced

little mutagenesis above background, whereas pso-AGT30 induced mutations in 1-2% of the vector molecules. At the eight hour point significantly induced mutagenesis was seen with pso-AG10, while psoAGT30 produced mutations in 5% of the target genes. The level of mutagenesis produced by pso-AGT20 was intermediate between the two. These results help to define the requirements for efficient *in vivo* triplex formation and mutation targeting.

TABLE 3. Targeted mutagenesis of pSupFG1 with COS cells

Oligo	K_d	Conc. μM	Time of UVA irradiation (hrs)	Mutation frequency	Mutants per total colonies
None	n.a.	n.a.	n.a.	0.03	7/7,600
pso-AG10	8×10^{-7} M	0.9	2	0.1	8/8,674
		0.2	8	0.1	3/3,148
pso-AGT20	7×10^{-9} M	0.45	2	0.1	18,00
		2.0	8	0.1	1/1073
pso-AGT30	3×10^{-9} M	0.45	2	0.3	/5,175
		2.0	8	0.5	3/651

Time course of UVA irradiation. The kinetics of intracellular targeted mutagenesis mediated by the mutagenic oligonucleotide was investigated. This process depends on the entry of the oligonucleotides into the cell, migration into the nucleus, and specific binding to the triplex target site. In order to achieve target site mutagenesis via site-specific generation of a psoralen photoadduct, these steps must occur by the time UVA irradiation is given. To determine the optimal time for UVA irradiation, a time course experiment was carried out in which the time of irradiation following pso-AGT30 addition to the

cell culture medium was varied. The results are shown in Table 4. As the time allowed for intracellular triplex formation before UVA irradiation was increased, the yield of targeted mutations also increased, with a frequency as high as 5.3% for the 8 hour point. Given sufficient time, therefore, the psoralen-linked oligonucleotides can enter cells, form a site-specific triple helix, and initiate targeted mutagenesis of an SV40 vector at frequencies in the range of 5%.

At the later time points, it is theoretically possible that lesions are targeted to vectors that are not subsequently replicated, or repaired before the molecules are transformed into bacteria for analysis. Hence, bacterial processing of persistent lesions produced in the COS cells may be occurring. Such vector molecules could escape the *Dpn* I restriction step designed to eliminate input vector molecules that had not replicated in the COS cells if they had already undergone replication prior to the triplex formation and UVA irradiation. However, to address this possibility, we allowed pso-AGT30 and pSupFG1a to form triplex *in vitro*, irradiated the complex with 1.2 J/m^2 of UVA irradiation, and used the sample to directly transform *E. coli* without passage through COS cells. A mutation frequency of .03% (4/15,000) was seen. Therefore the targeted mutagenesis observed in this experiment cannot be accounted for by the rescue of unprocessed vector molecules form the COS cells.

TABLE 4. Time dependence of targeted mutagenesis within COS cells

Oligonucleotide[a]	Time of UVA after oligo addition (hrs)	Mutation frequency %	Mutants/total
None	n.a.	0.03	4/14,700
pso-AGT30	1	1.1	101/9,400
"	2	2.1	178/8,663
"	4	3.6	474/13,339
"	8	5.3	343/6,515

[a] The oligonucleotide concentration used was 2 μM for all time points.

Time course of oligonucleotide treatment. In the above experiments, the cells were exposed to the mutagenic oligonucleotide within one hour after introduction of the SV40 vector DNA by electroporation. Although transfected SV40 vectors become covered in chromatin upon introduction into monkey cells, this process could be incomplete at such an early time point. An experiment was performed in which the cells were transfected with pSupFG1a, allowed to attach for 4 hours in growth medium, and incubated for 12 hours. The cells were then detached by trypsinization, washed three times in growth medium and incubated in the presence of 2 μM pso-AGT30 for two hours before irradiation with UVA. In this experiment, mutations were generated in the *supFG1a* gene at a frequency of 1.5% (150/10,175), in the same range as the frequency at the early time point for the same concentration of oligonucleotide (2.1 %). Here, even with more than enough time allowed for chromatin assembly on the SV40 vector DNA, the psoralen-conjugated oligonucleotide can still generate targeted mutations in the *supFG1a* gene of the vector.

This provides evidence that t[illegible]p[illegible]x for[illegible]ion can occur with chromatin.

The detection of targete[illegible]ns this experiment also confirms th[illegible] mutagenesis that observed i[illegible] intracellular triplex forma[illegible]. A[illegible]ter [illegible]owth of the cells for 12 hours, try[illegible]t[illegible]on, [illegible]lution, and extensive washing, it w[illegible] [illegible]nl[illegible]y that any vector DNA would persis[illegible]e ex[illegible]ellular medium at the time of oligo[illegible]e [illegible]tment.

Concentration dependence. [illegible]nc[illegible]ation dependence for targeted mut[illegible] wa[illegible]so investigated. The results [illegible] sh[illegible] i[illegible]ble 5 below. Following electropo[illegible] o[illegible] t[illegible] OS cells with the SV40 vector DNA, t[illegible] e s we[illegible]ncubated in the presence of pso-AGT3[illegible] a[illegible] [illegible]ions from 0.1 nM to 2 μM. UVA irradiation w[illegible] g[illegible] eight hours later. As can be seen, [illegible] but [illegible]tectable frequency of mutagenesis wa[illegible] o[illegible] [illegible] the extracellular oligonucleotide [illegible] e[illegible] tr[illegible]n in the nanomolar range. This is [illegible] s[illegible] w the K_d for triplex formation for pso-A[illegible] su[illegible] a being 3 x 10^{-9} M. Although intracellular [illegible]ligo[illegible]leotide concentrations were not dir[illegible]ct[illegible]y [illegible] su[illegible] in these experiments, other studies av[illegible]rt[illegible] hat treatment of mammalian cell[illegible] w t[illegible] ig[illegible]leotides produces concentrations wit[illegible] c[illegible] t[illegible] are in the same range as the give[illegible] e lu[illegible] concentrations. Significan[illegible] s of[illegible]geted mutagenesis were seen at an extra[illegible] lu[illegible] mutagenic

oligonucleotide concentrati n uM, sistent
with the data presented in a 3 .

TABLE 5. Concentration dependence of e t l m t sis within COS cells

Oligonucleotide	Concentration (nM) (hrs)	M	requ	Mutants/total
None	0	).(		2/3,400
pso-AGT30	0.1	).3		6/2,050
"	1	.(		12/2,200
"	10	.5		20/2,550
"	100	.7		19/2,860
"	1000	.		198/8771
"	2000	.		343/6515

[a] The values represent the frequency of u dic he oligonucleotide concentrations listed with T A ior eight hours after oligonucleotide addition to the cells.

Sequence analysis. A s l 36 tions induced in the *supF* gene b p T3(he 2 hour and 8 hour time points were an ed A sequencing. The results a s i 10. Of these, 28 (78%) were T:A t A nsv ons at bp 166 (the predicted psoralen i lat site) while 6 (17%) were deletion c riou tes spanning the triplex targe s . The erved frequency of deletions may e nd mate, since larger deletions inv vi equ surrounding the *supFGla* ge w ra ate essential genes for vector e ion herefore, molecules undergoing such s wc not be detected in the assay. Th e lts nstrate the specificity of the int c ar ion targeting and suggest that n ra ducts can

generate either specific po:nt ... or deletions at target sit .

Specificity: the effect of ... ology for triplex formation. The specif ... y ... e mutagenic oligonucleotide-mediat ... ed mutagenesis may be influenced by sev ... factors, one of which is the existen ... lter ... sites in the DNA having partial homo ... x formation by the psoralen-co ... d ... enic oligonucleotide. This issu ... by comparing targeted mutagenesis ... he ... *G1a* and *supFG2* genes by pso-AGT43. Pso-AGT4 ... designed to form a triple helix at ba ... 209 of *supFG2*. In the anti-parall ... tches exactly the 43 bp site in su ... r the two T interruptions at base pair ... (95% homology). In contrast, pso-... y 65% homology for triplex formation ... a. The first 30 nucleotides in pso-AGT4 ... xactly those in pso-AGT30, and so ... 8 out of 43 nucleotide homology for ... on with *supFG1a* (taking into account ... le mismatches at 180 and 183). ... GT30, the tethered psoralen in pso-AGT4 ... to intercalate at base pairs 166-1 ... ceptually similar experiments, a compa ... GT30 in targeting both the *supFG1a* ... it is designed to form triplex) a ... *supF* gene (to which it has only ... cleotide homology for triplex formati ... In these experiments, the electroporate ... were incubated in 2 μM concentrat ... tagenic oligonucleotide for two hour ... irradiation was given.

Pso-AGT43 effectively ta ... s to *supFG1a* as shown below in Ta ..., pso-AGT43 failed to induce ... s ... e *supFG1a*

gene above the background f[illegible] gue[illegible], [illegible]esting that the partial homology f[illegible] tr[illegible]p[illegible]ex [illegible]mation with *supFG1a* is insufficien[illegible] [illegible]ediat[illegible] significant *in vivo* interactions. In t[illegible] [illegible] wa[illegible] p[illegible]so-AGT30 can target mutations to *supi*[illegible] t[illegible] not effective in inducing mutat[illegible] t[illegible] [illegible]odified *supF*. These results demons[illegible]ate [illegible]ne [illegible]ficity of the intracellular targeting

The frequency of targete[illegible] [illegible]ions [illegible]duced by pso-AGT43 in *supFG2* was not [illegible] r th[illegible] that seen with pso-AGT30 and *supFG1a*, [illegible] r[illegible] [illegible] the increase observed in going [illegible] - [illegible]o pso-AGT20 to pso-AGT-30 (Ta[illegible]e [illegible]. I[illegible] possible that the 43 nucleotide long -ri[illegible] o[illegible] gc [illegible]cleotide (pso-AGT43) may be subject [illegible] [illegible]iven self-association to form G-[illegible]ts, [illegible]ting its effectiveness within the ce[illegible]

TABLE 6. Targ[illegible] [illegible]fic[illegible]

Oligonucleotide	Target gene	Homolog[illegible] for triple[illegible] formation[illegible]	[illegible]% [illegible]tior [illegible]nc[illegible] [illegible]%	[illegible]tants/tot
pso-ATG43	*supFG1a*	30/43	[illegible]	[illegible] 6887
pso-AGT43	*supFG2*	41/43		[illegible] 3/16,425
pso-AGT30	*supFG1a*	28/30		8/8663
pso-AGT30	*supF*	17/30	[illegible]2	[illegible] 300
pso-TCGA30	*supFG1a*	12/30		3,740

[a] The values represent the fraction of b[illegible] [illegible]s in t[illegible] [illegible]-parallel triple helix motif (G for G:C bp and either [illegible] A:[illegible]) between the listed oligonucleotide and the designated ta[illegible]

[b] The values represent the frequency of [illegible] [illegible] served using oligonucleotide concentrations of 2 μM an[illegible] [illegible] i[illegible] [illegible] given two hours after oligonucleotide addition to the [illegible]s.

The foregoing results d[illegible] [illegible]e [illegible]s[illegible] of a triple helix forming oligor[illegible] [illegible]le [illegible] [illegible]nerate

targeted mutations within m c Linkage
of the oligonucleotide to a len,
confers sequence specificit of the
psoralen, which is delivere ed site
via intracellular triple he ac An
oligomer with a K_d in the r 0- better
is preferred for significar el
interactions. Other factor ic fect
triplex formation under phy ons,
such as oligonucleotide bas id
backbone structure, could be this type
of assay. Because selected es w in and
adjacent to the *supF* gene c mi d
without disrupting suppress r vit ditional
target sites designed to sp ly ne other
aspects of intracellular tr r could be
constructed.

Sequence analysis of the t ted by
pso-AGT30 revealed a high s ty f :A to
A:T transversions at the pr ps
intercalation site (bp 166), es of the
time course experiment ind t ocesses
of oligonucleotide entry in
intracellular formation of e occur
over several hours. The i y of
mutations seen at the later , in
part, reflect the time need ne cesses
as well as the interplay o ar ication
in forming the targeted DN tations.

Modifications and var t ent
invention, mutagenic triple
oligonucleotides, as well a
thereof, will be obvious to n the
art from the foregoing deta . Such
modifications and var o come
within the scope of the appe

We claim:

1. A mutagenic oligonuc f e-directed mutagenesis of a r nucleic acid molecule comprising l rated into a single-stranded ol g ing a sequence that forms a tripl eic acid molecule with a target regio e e-stranded nucleic acid molec

2. The mutagenic oligon e aim 1 wherein the mutagen is sele m oup consisting of psoralen, a r a alkylating agent, a cis-p a hematoporphyrin, a hemato o d ve, mitomycin C, a radionuclide m e that interacts with radiation to m ic.

3. The mutagenic olicor e aim 1 wherein the mutagen cause: o. e double-stranded nucleic a i l e presence of light.

4. The mutagenic oligon aim 3 wherein the mutagenic chemic methyl-4,5',8-trimethylpsoralen.

5. The mutagenic olicon e im 1 wherein the oligonucleoti e between 20 and 30 nucleotide base

6. A method for site-d is of a nucleic acid molecule compr of:

a) hybridizing a m cleotide to a target region of a dou leic acid molecule, wherein th oligonucleotide comprises a rated into a single-stranded n rms a triple-stranded nucleic a he target region; and

b) mutating the do cleic acid molecule.

7. The method of cla[illegible] additional step of activ[illegible] mut[illegible] prior to the mutation step.

8. The method of cla[illegible] in t[illegible] itagen is selected from the group [illegible] of [illegible]len and acridine orange and is a[illegible]y li[illegible]

9. The method of claim [illegible]he in t[illegible] itagen is selected from the group [illegible] of [illegible]line orange, an alkylating ag[illegible] a [illegible]-pl[illegible] analog, a hematoporphyrin, a hemat[illegible] in de[illegible]tive, mitomycin C, a radionucl[illegible] mol[illegible] that interacts with radiation [illegible] mut[illegible]c.

10. The method of cla[illegible] w[illegible] ei[illegible] [illegible]utation alters the activity of t[illegible] [illegible] str[illegible] nucleic acid molecule.

11. The method of cl[illegible] w[illegible] ein[illegible] [illegible]ouble-stranded nucleic acid mo[illegible] [illegible] a ge[illegible]

12. The method of cl[illegible] ein[illegible] [illegible]gene is an oncogene.

13. The method of cla[illegible] w[illegible] ei[illegible] [illegible]ene is a defective gene.

14. The method of cl[illegible] 3 [illegible]ein[illegible] gene is a defective human ß-hemo[illegible]r[illegible]e.

15. The method of cl[illegible] w[illegible] ein[illegible] [illegible]ouble-stranded nucleic acid mo[illegible] al[illegible] portion of a viral genome.

16. A method of prod[illegible] a [illegible]ta[illegible] oligonucleotide comprisi[illegible] [illegible]ps [illegible]

a) synthesizing a [illegible]i[illegible] [illegible]cl[illegible]e substantially complement[illegible] [illegible]on t[illegible] [illegible]ird strand binding code to a [illegible] [illegible]gi[illegible] [illegible] double-stranded nucleic acid mo[illegible] [illegible]d

b) incorporatin[illegible] [illegible] [illegible]r oligonucleotide.

17. The method of cl[illegible] [illegible]ce[illegible] mutagen is covalently linked to [illegible] [illegible]d[illegible].

18. The method of cla[illegible]e mutagen is incorporated into the o[illegible]ring synthesis of the oligonucl[illegible]d.

19. The method of cl[illegible]ein mutagen is bound to the oligonuc[illegible] ph[illegible]ivation.

20. The method of clai[illegible]herein mutagen is selected from the grou[illegible]inc soralen, acridine orange, an alkyl[illegible]nt [illegible]- platinum analog, a hematop[illegible]y[illegible], a hematoporphyrin derivative[illegible]t [illegible]in radionuclide, and a molecu[illegible]nt[illegible] with radiation to become mutage[illegible].

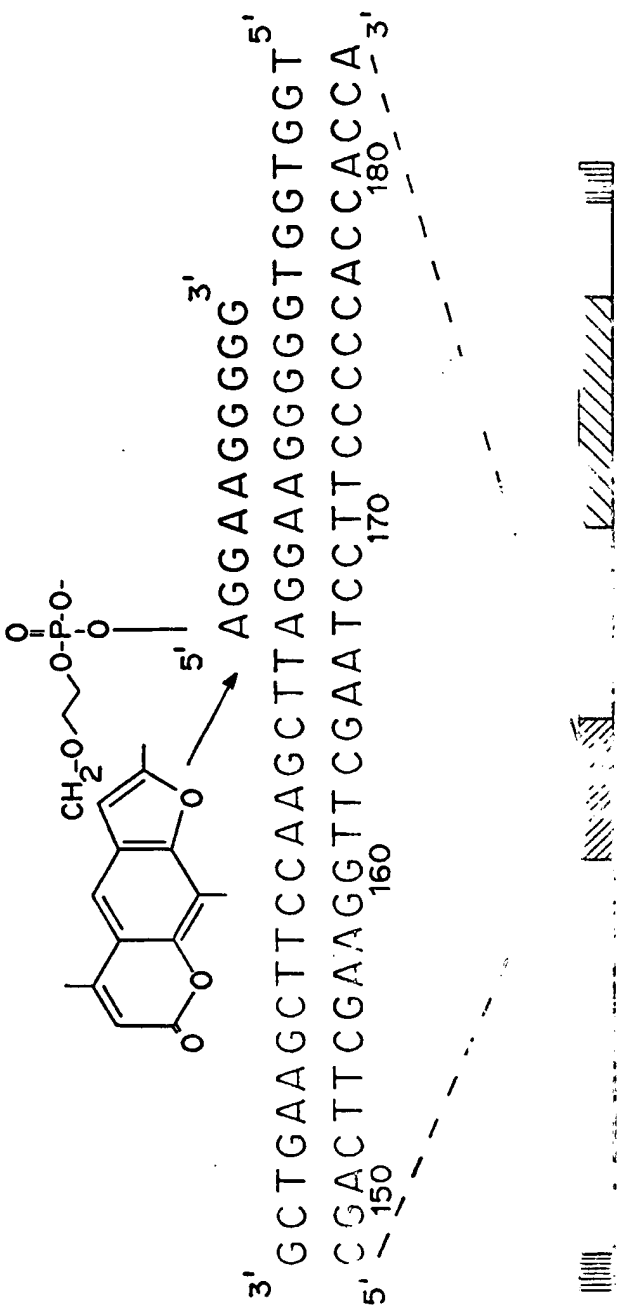

3'
GCTGAAGCTTCCAAGCTTAGGAAGGGGGTGGTGGT
CGACTTCGAAGGTTCGAATCCTTCCCCCACCACCA
5'
150
160
170
180
3'
5'
AGGGAAGGGGG
3'
5'
O=P-O-
O
CH2-O
O
O
O

FIG. 1

3'
TAAACTATACTACGCGGGG
40 50
2/11
90 100 110 120 130 140 150 160 * 170 180
GTAAAAGCATTACCTGTGGTGGGGTTCCCGAGCGGCCAAAGGGA GCAGACTCTAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACCACCA
+ ++ ++ +++ +++++ ++ ++++ ++++ ++ +++ +++++ +++++ ++ ++ ++++ +++++ ++ +++++ ++++++ +++++ +
pre-tRNA (58-98) Suppressor tRNA (99-183)
3'

FIG. 2a

FIG. 2b

A
G TC
3' G TC AT
TAAACTATACTACGCGGGG
40 50
ATTGATATGATGCGCCCC
5'
100
120
140
150
170

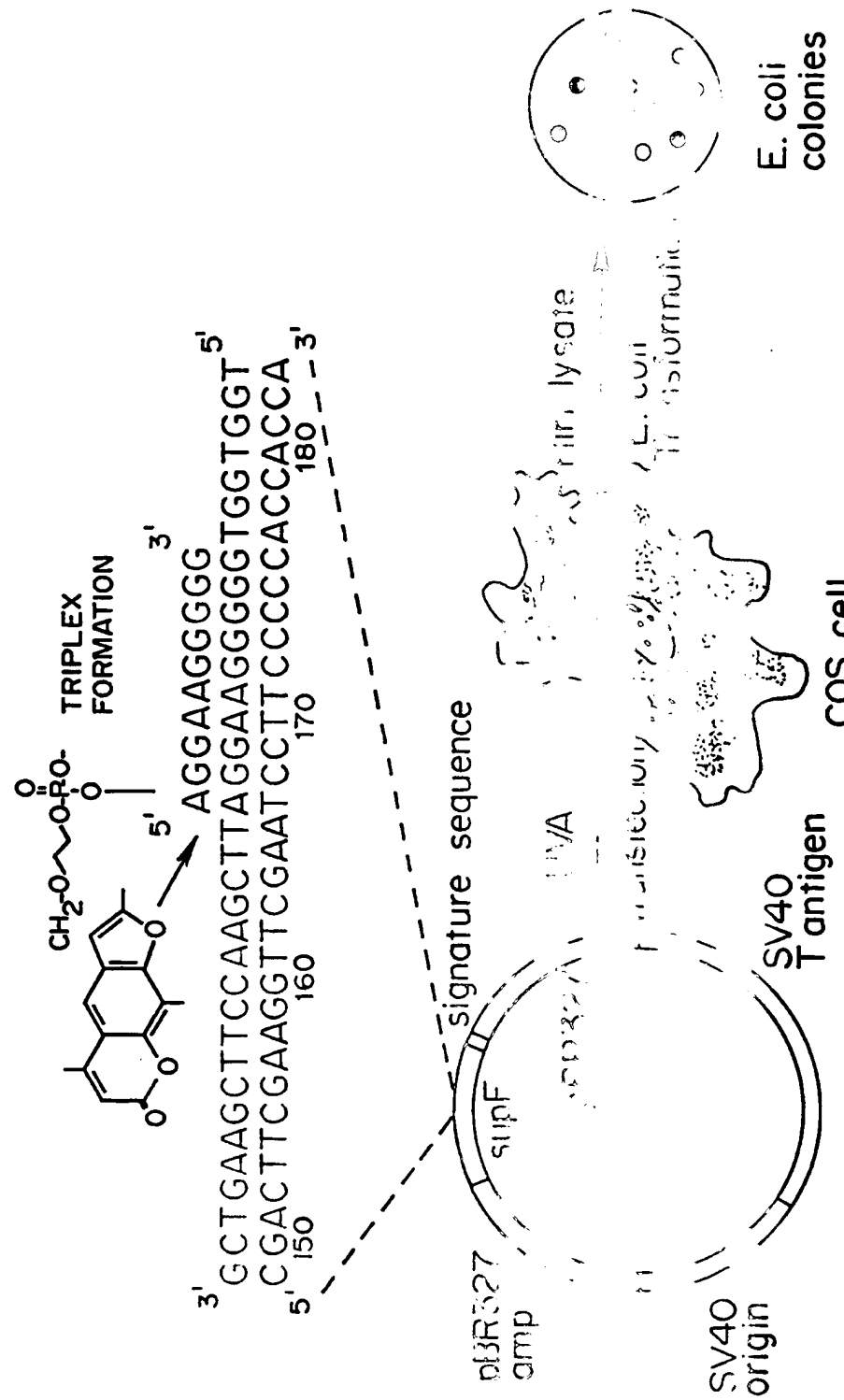

TRIPLEX FORMATION
5'
3'
AGGAAGGGGG
GCTGAAGCTTCCAAGCTTAGGAAGGGGGGTGGTGGT
CGACTTCGAAGGTTCGAATCCTTCCCCCACCACCA
3'
5'
150
160
170
180
3'
5'
signature sequence
pBR327 amp
SV40 origin
SV40 T antigen
Shuttle vector pSP189
COS cell
E. coli colonies

FIG. 3

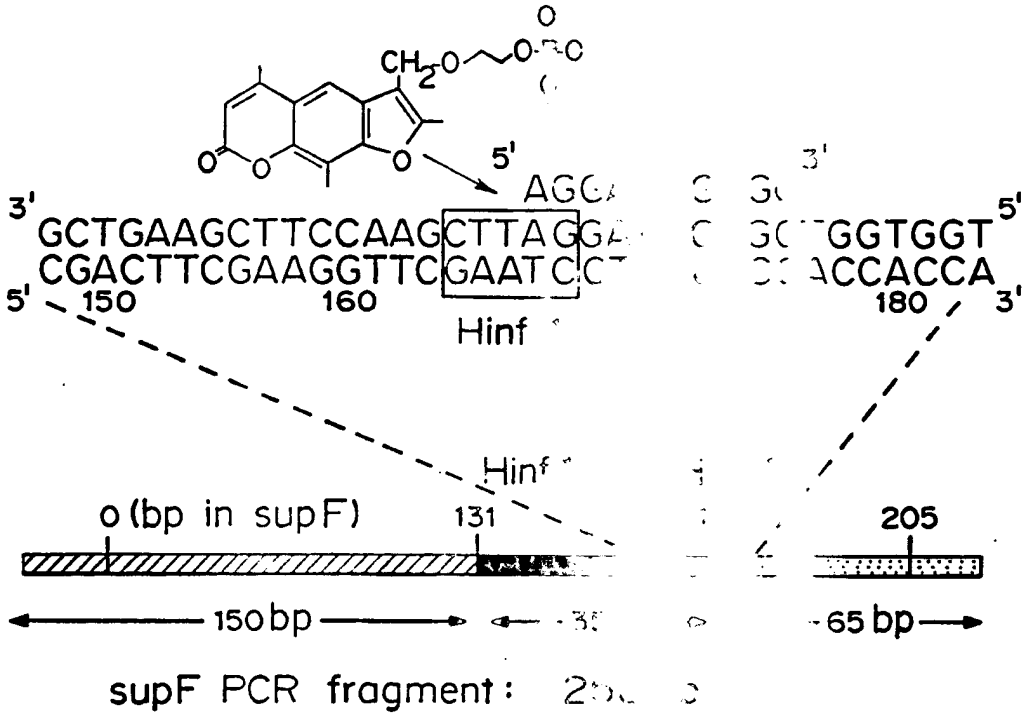

5'
3'
GCTGAAGCTTCCAAGCTTAGG
CGACTTCGAAGGTTCGAATCC
ACCACCA
GGTGGT
5'
150
160
180
3'
Hinf
Hinf
0 (bp in supF)
131
205
150 bp
65 bp
supF PCR fragment:

FIG. 4

FIG. 5a

SUBSTITUTE S

FIG. 5b

A
G TC
3' G TC AT
TAAACTATACTACGCGGGG
40 50
ATTTGATATGATGCGCCCC
5'
T
T
TC
A
A
A
A
A
A
AG C AA C G 5'
3'

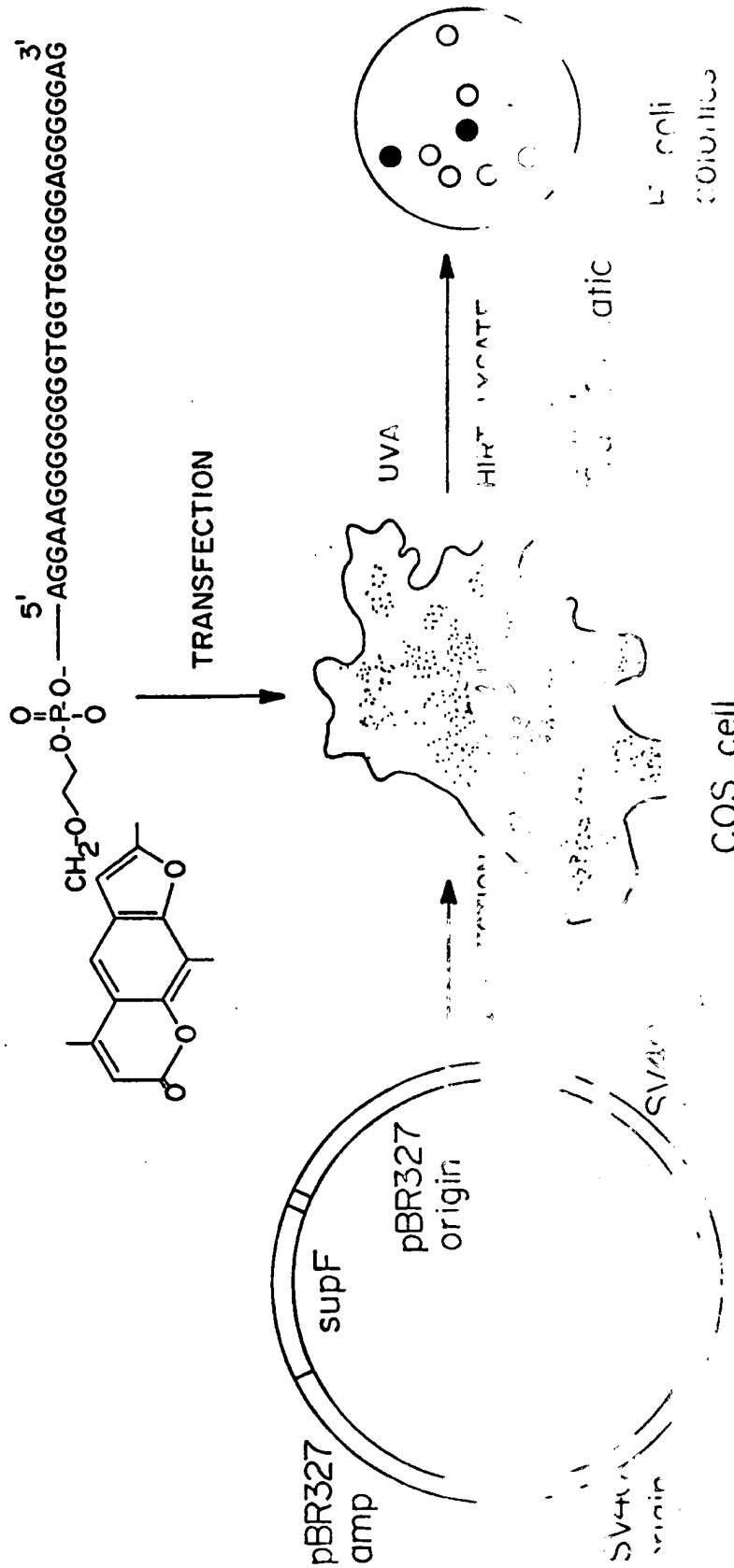
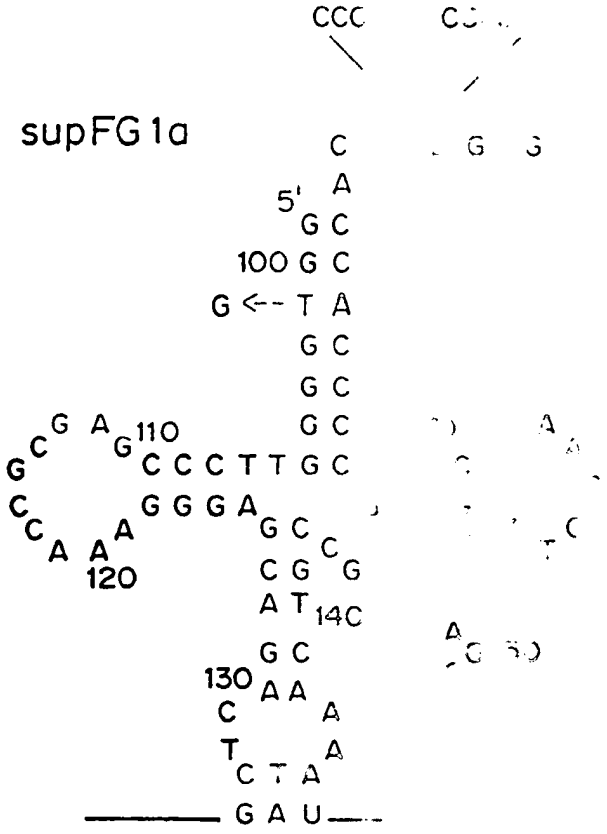

5'
AGGAAGGGGGGGGTGGTGGGGGAGGGGGAG
3'
TRANSFECTION
UVA
pBR327
amp
supF
pBR327
origin
Shuttle vector
pSupFG1a
COS cell

FIG. 6

FIG.



Predicted triple helix:

pso-AGGAAGGGGGGGGTGGT(... GAG 3' pso-AGT30
3' AGGAAGGGGGGGGTGGT(... GAG 5' *supFG1a*
5' TCCTTCCCCCCCCACCA(... CTC 3' duplex
167 180

FIG. 8

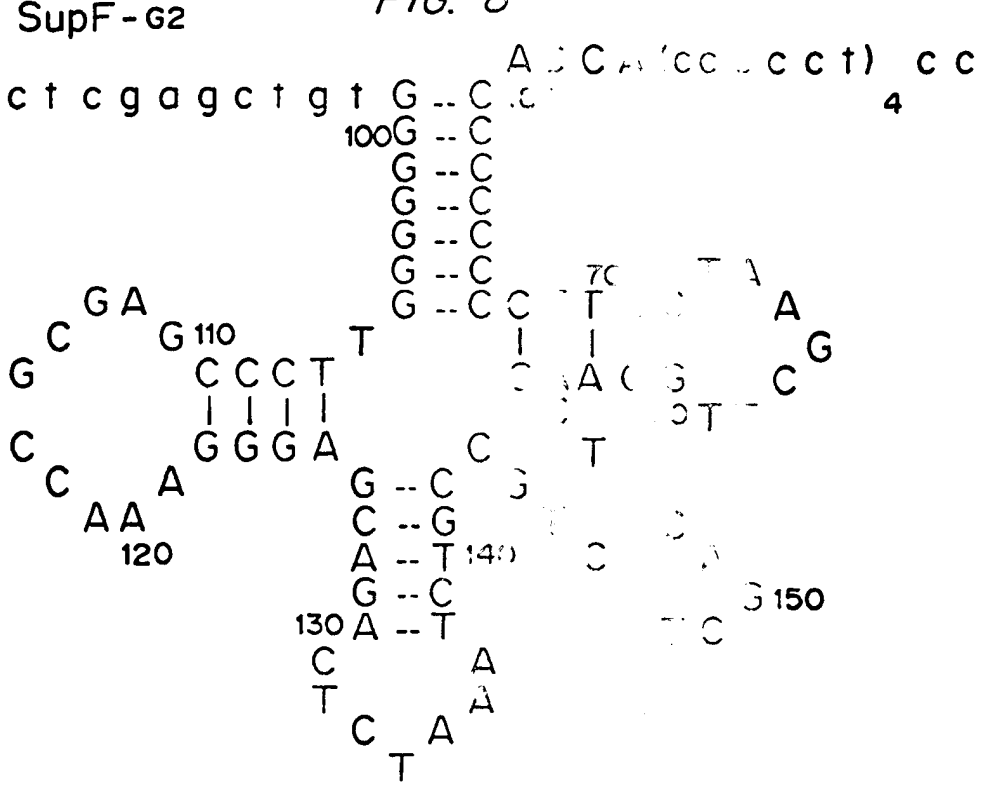


FIG. 9

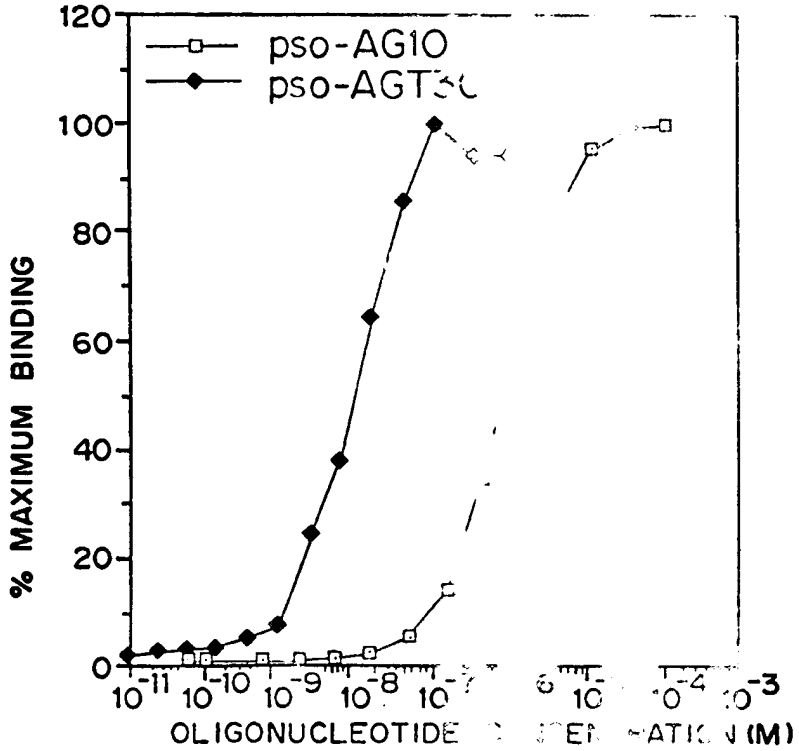


FIG. 10

SUBSTITUTE S